

Designing for Efficient Socialization and Ideological Critique

Aaron Halfaker

ahalfaker@gmail.com

Wikimedia Foundation

R. Stuart Geiger

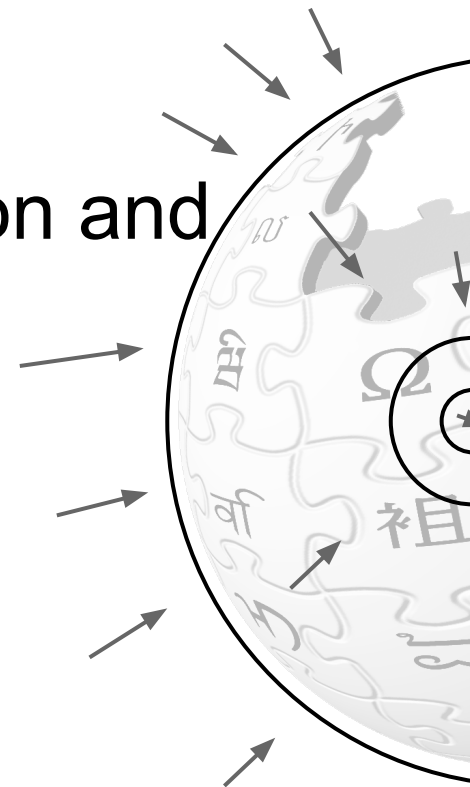
stuart@stuartgeiger.com

UC Berkeley iSchool

Loren Terveen

terveen@cs.umn.edu

University of Minnesota



Hello & welcome!

Hello & welcome!

Name: Aaron Halfaker

Hello & welcome!

Name: Aaron Halfaker

Occupation: Research Scientist @ Wikimedia

Hello & welcome!

Name: Aaron Halfaker

Occupation: Research Scientist @ Wikimedia



WIKIPEDIA
The Free Encyclopedia



WIKIMEDIA
COMMONS

a multilingual tree
encyclopedia

Wiktionary

['wɪkʃənri] *n.*,
a wiki-based Open
Content dictionary

וילון ['wɪlɔn kɔʔɪ]

Hello & welcome!

Name: Aaron Halfaker

Occupation: PhD Candidate @ UMN

grouplens

UNIVERSITY OF MINNESOTA



- Complex social problem

- Complex social problem
- Interesting strategy for a social solution

- Complex social problem
- Interesting strategy for a social solution
 - Challenges ideologies
 - Powerful new tool for Wikipedians

1. Wikipedia's success & socio-technical problems

1. Wikipedia's success & socio-technical problems

2.

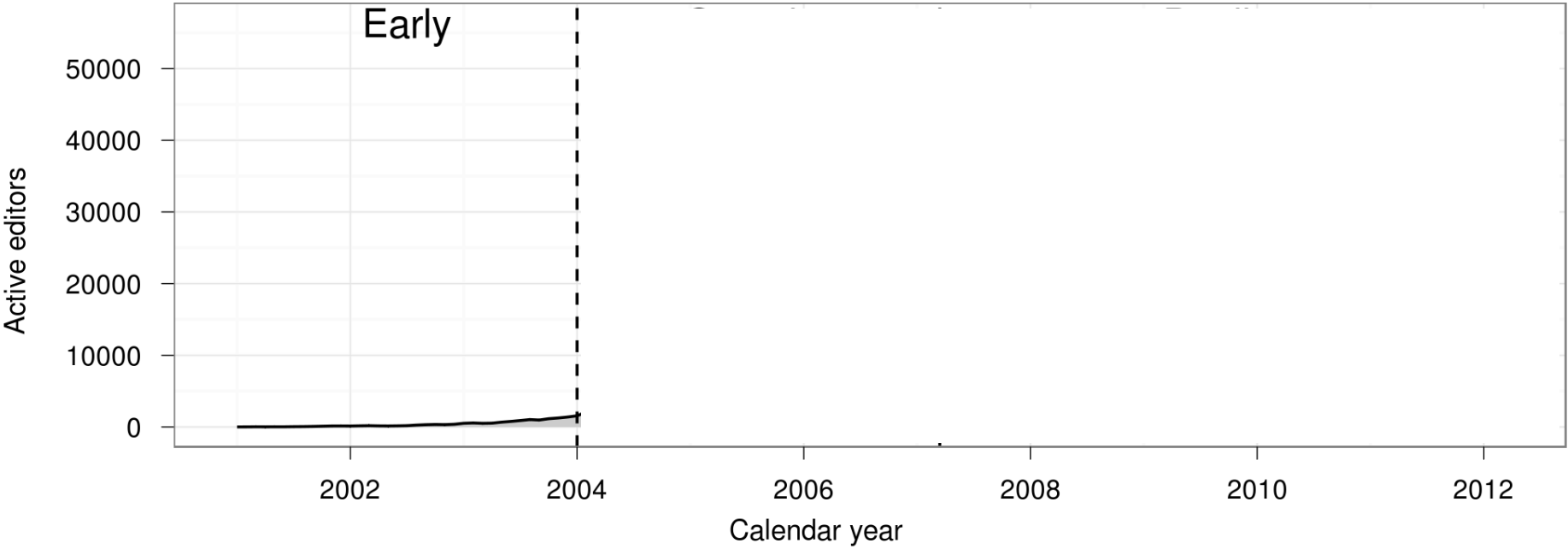


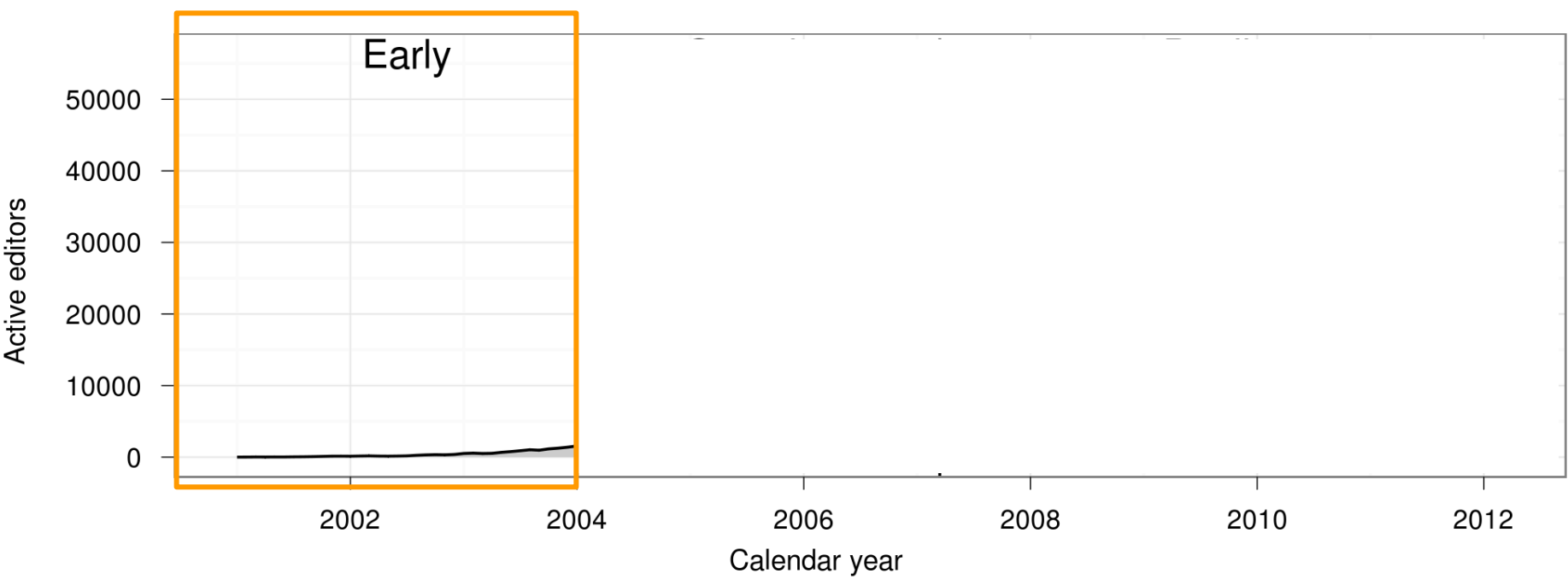
Power socio-technical tool

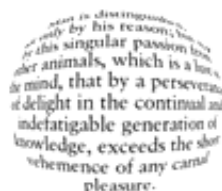
Means of ideological critique

Part 1.

Wikipedia's socio-technical problems







WIKIPEDIA
The Free Encyclopedia

[Main Page](#)
[Recent changes](#)
[Random page](#)
[Current events](#)

[Edit this page](#)
[Discuss this page](#)
[Older versions](#)
[What links here](#)
[Related changes](#)

[Special pages](#)
[Bug reports](#)

[Main Page](#) | [Recent changes](#) | [Edit this page](#) | [Older](#)

[versions](#) |

[Printable version](#)

209.237.238.163

[Log in](#) | [Help](#)

Homepage

From Wikipedia, the free encyclopedia.

On the [World Wide Web](#), a **homepage** is the starting point, or root of a [website](#), a group of pages belonging to the same person, group, business or any other organization. The home page usually includes basic information on the theme of the whole [website](#), and links to other related pages.

[Wikipedia](#)'s homepage is the [Main Page](#) which outlines the project and gives links into the body of articles.

During the [1990s](#), a homepage often referred to someone's personal page, with information about themselves. This idea of a homepage is being superseded by the [blog](#).

Most [web browser](#) applications allow a particular [URL](#) to be specified as the homepage for the browser: the [web page](#) which is displayed when the application is started.

[Edit this page](#) | [Discuss this page](#) | [Older versions](#) | [What links here](#) | [Related changes](#)

[Main Page](#) | [About Wikipedia](#) | [Recent changes](#) |

This page has been accessed 397 times. It was last modified 09:16 Sep 24, 2002. All text is available under the terms of the [GNU Free Documentation License](#).



WIKIPEDIA
The Free Encyclopedia

[Main Page](#) | [Recent changes](#) | [View source](#) | [Page history](#)
[Printable version](#) | [Disclaimers](#)

Not logged in

[Log in](#) | [Help](#)

Main Page

From Wikipedia, the free encyclopedia.

Wikipedia is a [multilingual](#) project to create an accurate and comprehensive, [free-content](#) encyclopedia. We started in January 2001 and are currently working on 203744 [articles](#) in the English version. Visit the [help page](#) and experiment in the [sandbox](#) to learn how **you** can edit any article right now.

Selected Articles

In the news: [Peter Beattie](#) - [Hans Blix](#) - [Sri Lanka](#) - [Moscow Metro](#) - [Ricin incident](#)
Recent deaths: [Rebeca Martinez](#) - [Abdul Rahman Munif](#) - [Eleanor Holm](#) - [Carlie Bruscia](#)
New articles: [US Eighth Army](#) - [Trugemanna](#) - [William Jessop](#) - [Superb Fairy-wren](#)
Featured articles: [Paradox](#) - [Chess](#) - [Éire](#) - [Io \(moon\)](#) - [White Rose](#) - [Vampire](#) - [Punk rock](#)
Anniversaries: [February 9](#) - [The Beatles](#) on *The Ed Sullivan Show* - [Mandaluyong City](#) - [Hollywood Walk of Fame](#) - [Volleyball](#) - [Emperor Zeno](#) - *The Simpsons*

Encyclopedia

Mathematical and Natural Sciences

[Astronomy](#) - [Biology](#) - [Chemistry](#) - [Computer science](#) - [Earth science](#) - [Ecology](#) - [Health science](#) - [Mathematics](#) - [Physics](#) - [Statistics](#)

Applied Arts and Sciences

[Agriculture](#) - [Architecture](#) - [Business](#) - [Communication](#) - [Education](#) - [Engineering](#) - [Family and consumer science](#) - [Government](#) - [Law](#) - [Library and information science](#) - [Medicine](#) - [Politics](#) - [Public affairs](#) - [Software engineering](#) - [Technology](#) - [Transport](#)

Social Sciences and Philosophy

[Anthropology](#) - [Archaeology](#) - [Economics](#) - [Geography](#) - [History](#) - [History of science and technology](#) - [Language](#) - [Linguistics](#) - [Mythology](#) - [Philosophy](#) - [Political science](#) - [Psychology](#) - [Sociology](#)

Culture and Fine Arts

Community

Writing Articles

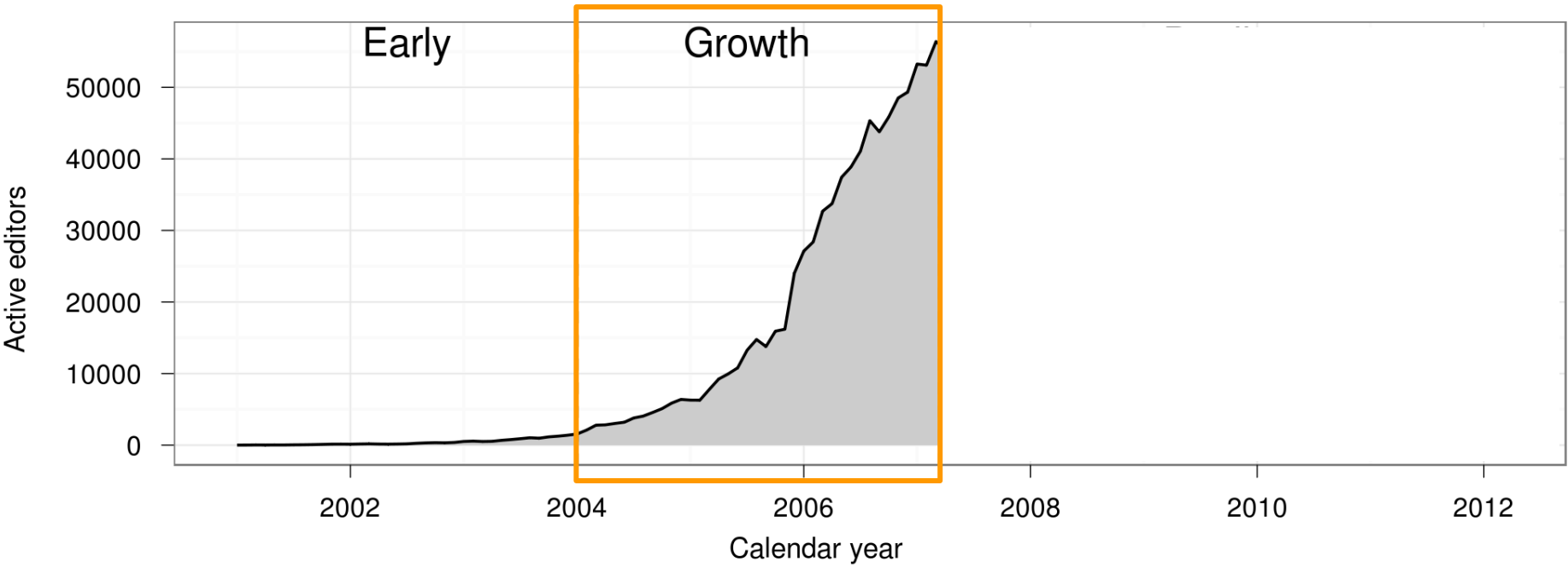
[Welcome, newcomers](#) - [Policy](#) - [Copyrights](#) - [How to edit](#) - [Guide to layout and style](#) - [WikiCivics](#) - [Public domain](#) and [shared resources](#) - [Requested articles and pictures](#) - [Featured articles and pictures](#) - [Reference desk](#) - [Projects](#)

About the Project

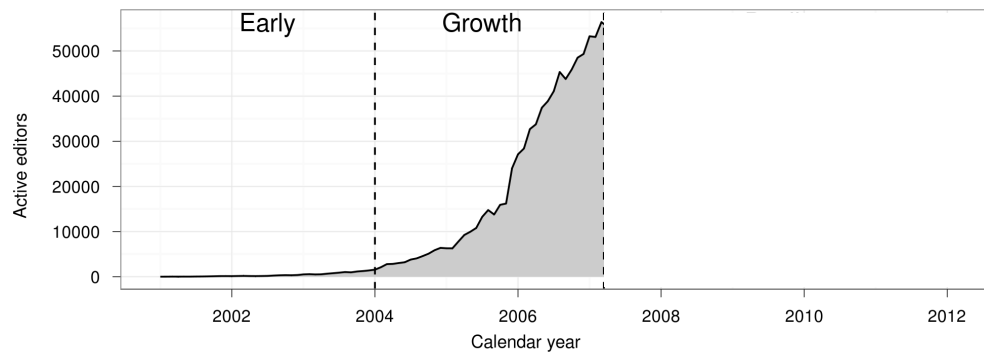
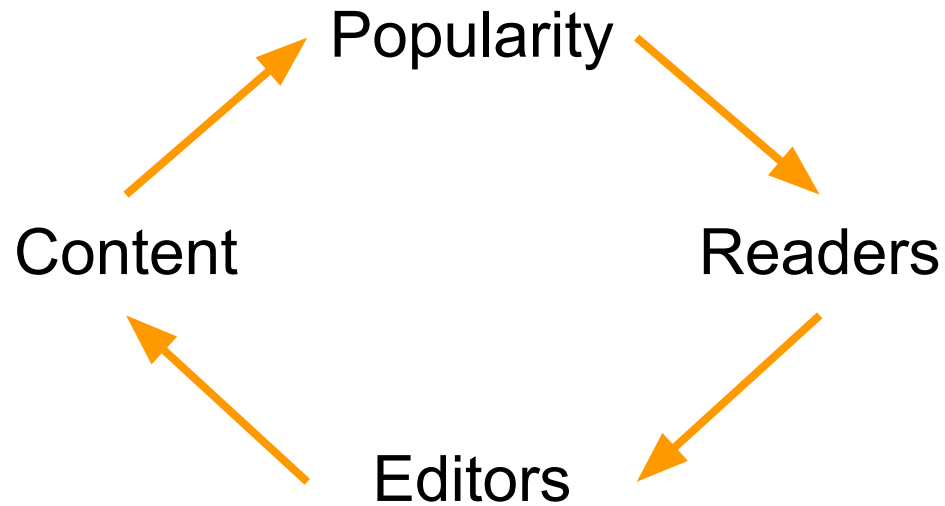
[About Wikipedia](#) - [Wikimedia](#) - [Announcements](#) - [FAQ](#) - [Contact us](#) - [Village pump](#) - [Chat](#) - [Mailing lists](#) - [Moderator BBS](#) - [Wikipedians](#) - [Friends of Wikipedia](#) - [Utilities](#) - [Software](#) - [Statistics](#) - [Donations](#) - [Download Wikipedia: SQL dumps](#) - [TomeRaider format](#) - [Wikipedia press coverage](#): [Google News](#) - [Yahoo! News](#)

In Other Languages

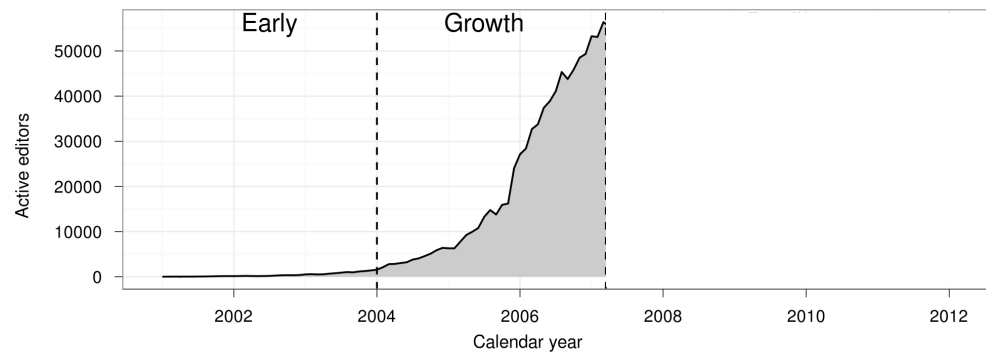
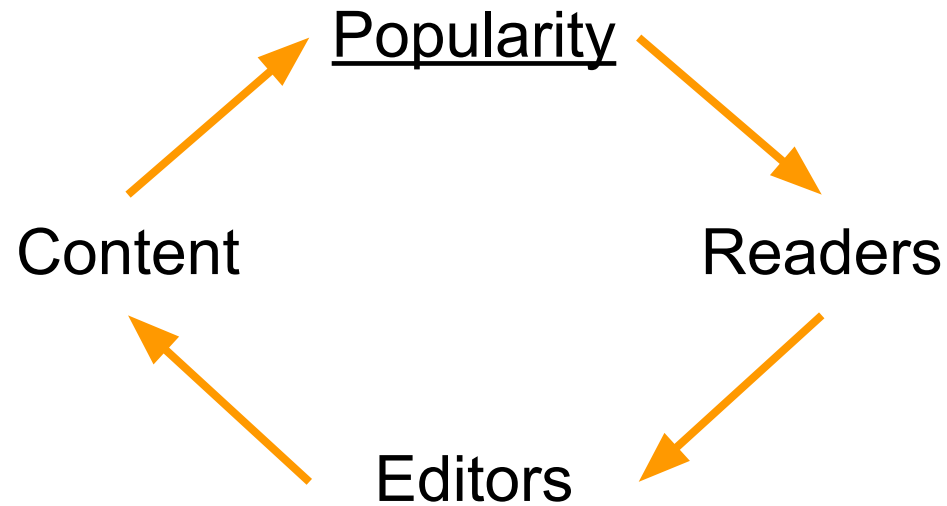
[Wikipedia language list](#) - [Afrikaans](#) - [العربية \(Araby\)](#) - [Bahasa Melayu](#) - [Bosanski](#) - [Български \(Bulgarian\)](#) - [Català](#) - [Česká](#) - [Corsu](#) - [Cymraeg](#) - [Dansk](#) -



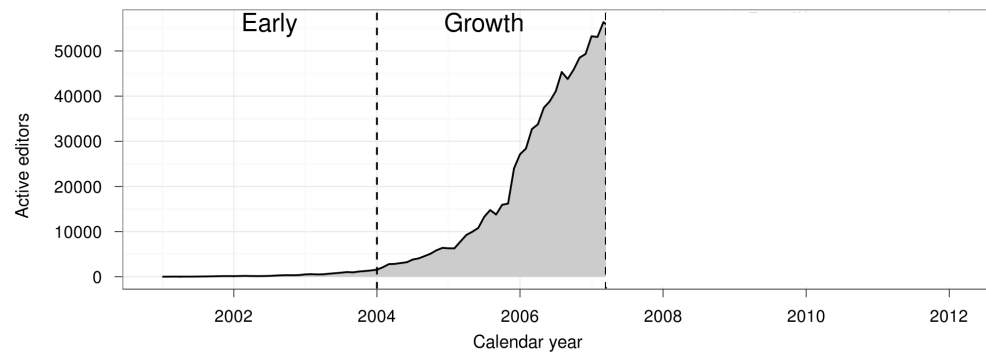
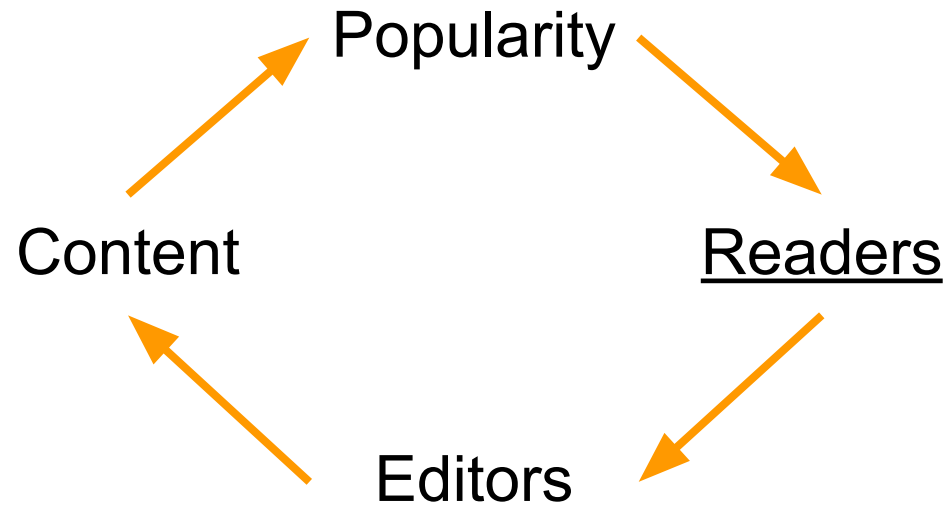
Wikipedia's Growth Model



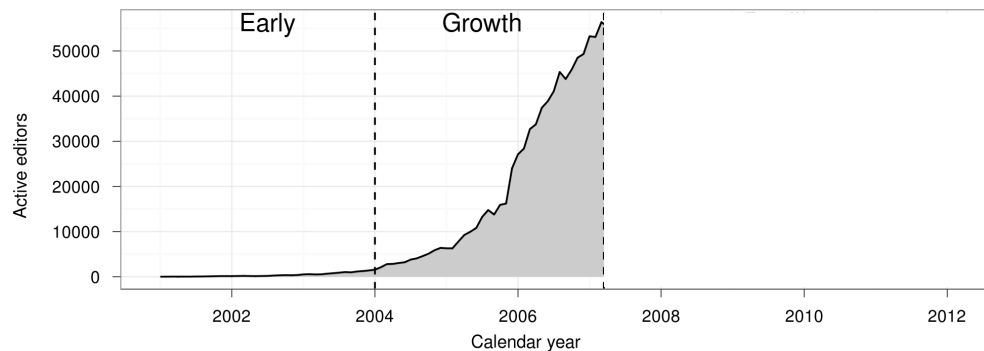
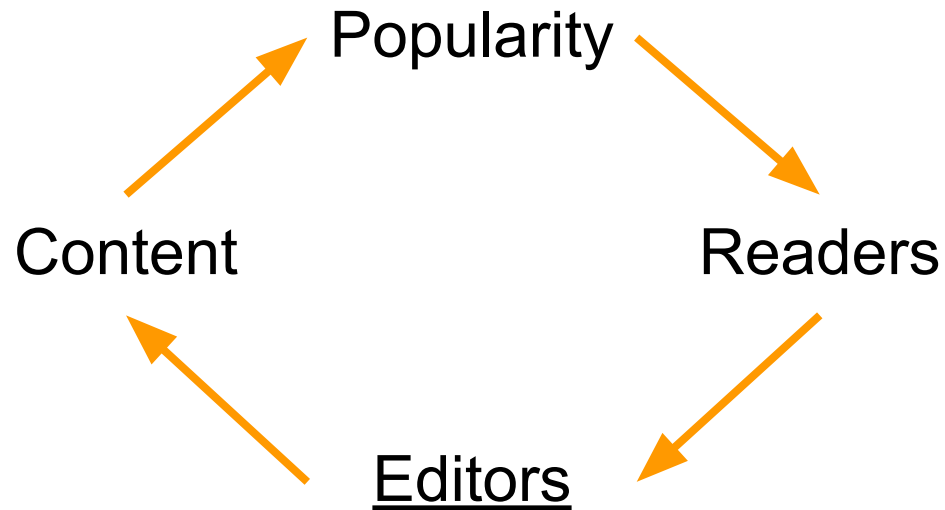
Wikipedia's Growth Model



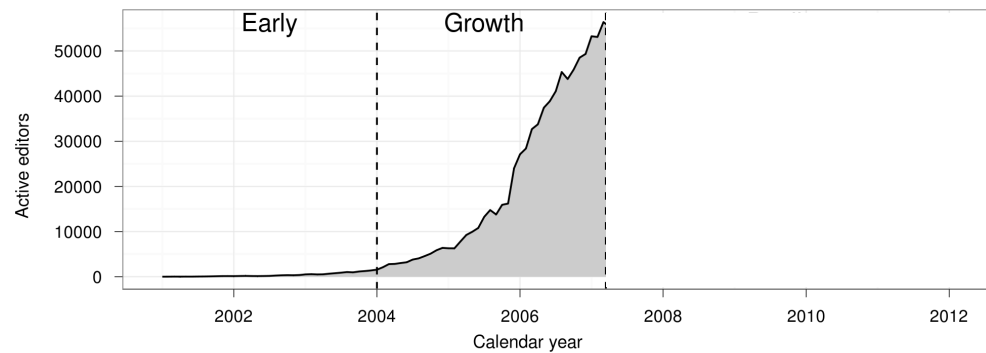
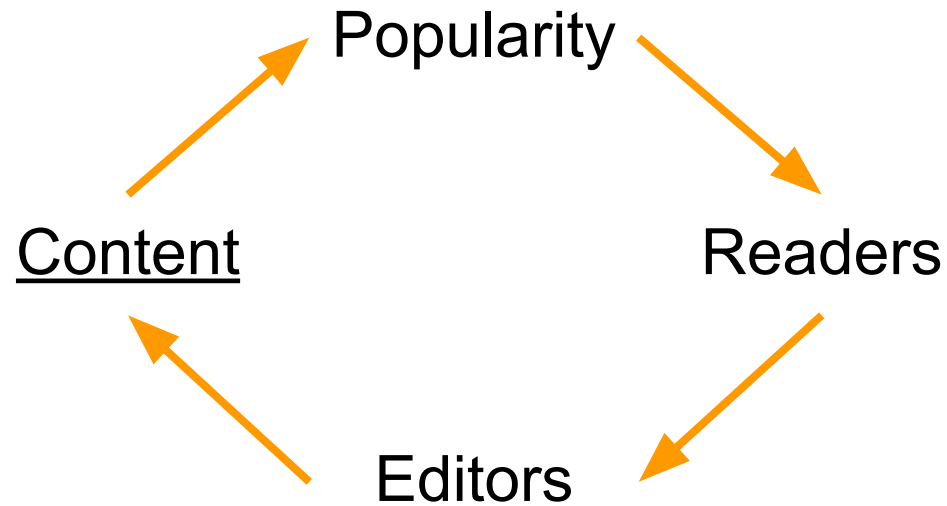
Wikipedia's Growth Model



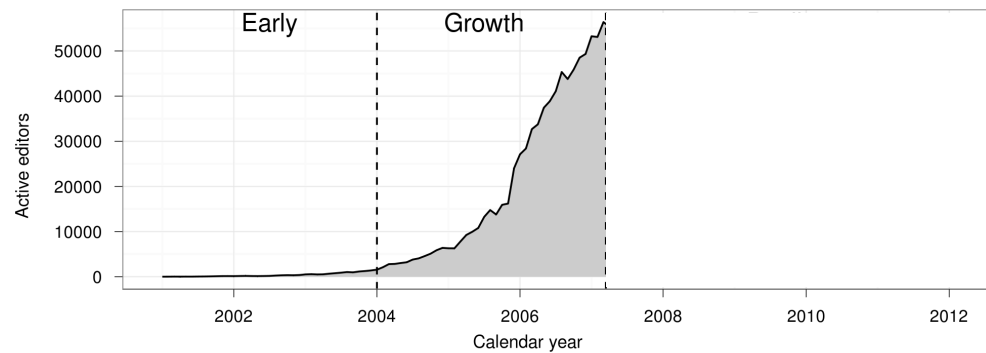
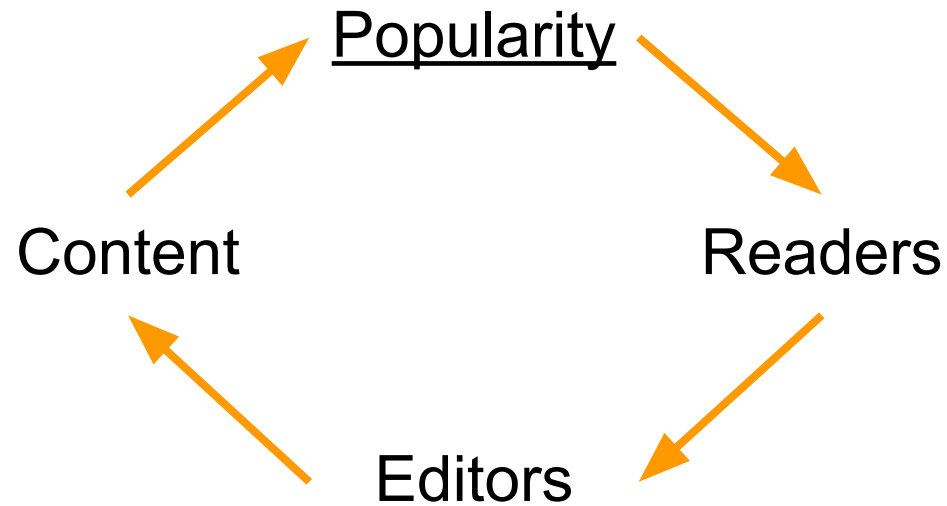
Wikipedia's Growth Model



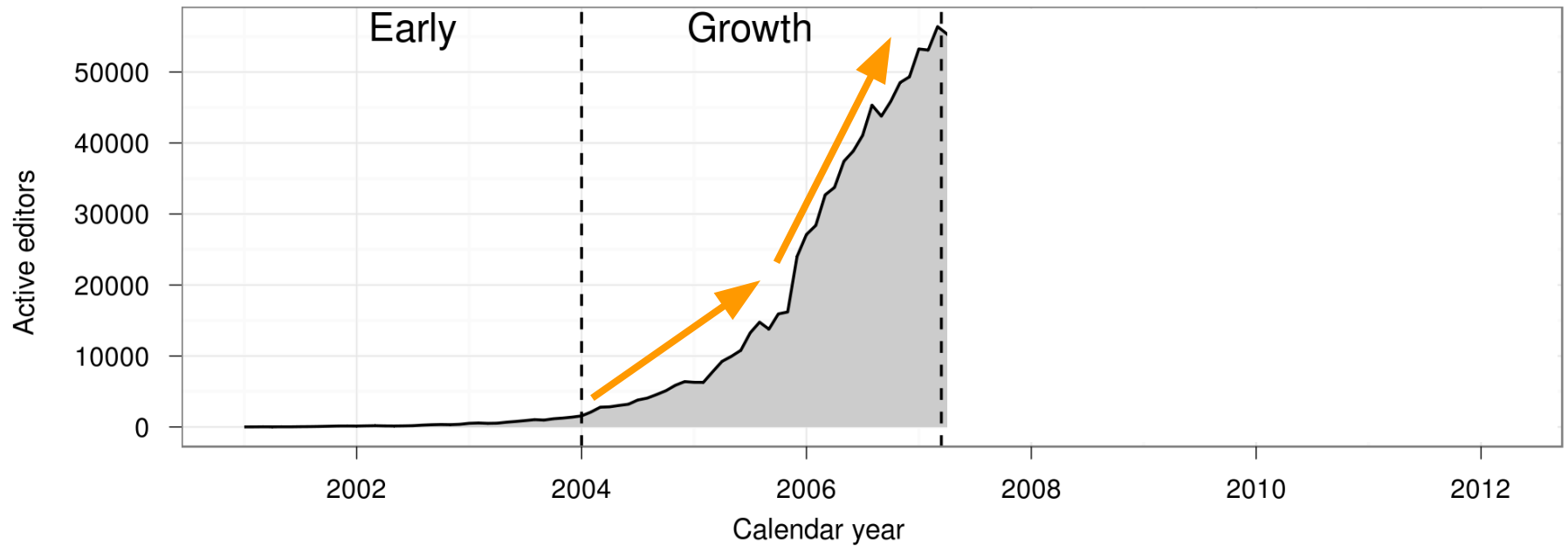
Wikipedia's Growth Model



Wikipedia's Growth Model



Wikipedia's Growth Model



Growth is Good, right?

Growth is Good, right?

- Popularity → Vandalism & Spam

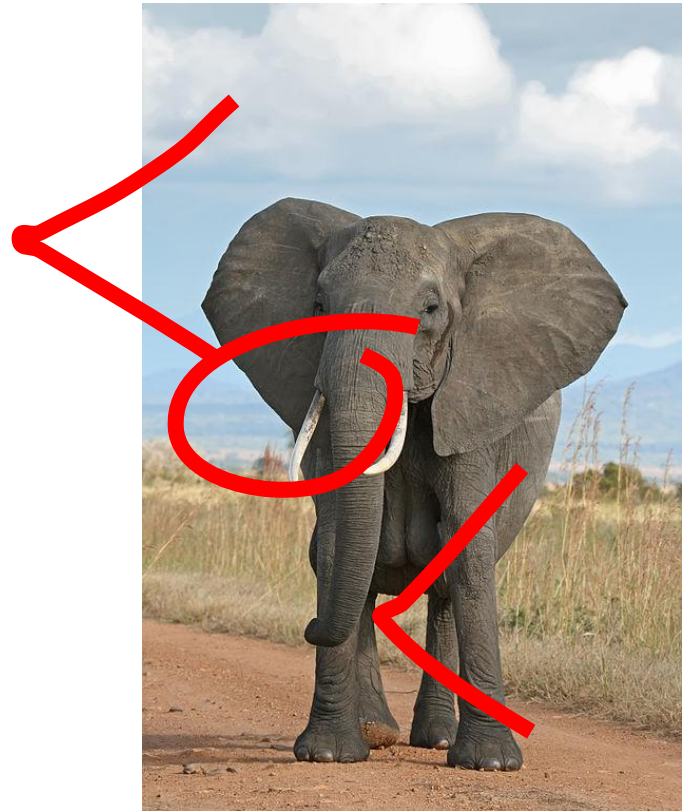
Growth is Good, right?

- Popularity → Vandalism & Spam



Growth is Good, right?

- Popularity → Vandalism & Spam



Growth is Good, right?

- Popularity → Vandalism & Spam



ARE YOU SMARTER THAN JUSTIN BIEBER?

His IQ is above **102!** Most People **CANNOT** beat him. Can you beat him?

[f Connect with Facebook](#)

👍 37,527 people like this. Be the first of your friends to like this.

Are You Smart Enough!?

Connect with your Facebook Account to see if you're SMARTER than Justin Bieber?

The image shows a promotional graphic for a Facebook link. At the top, the text 'ARE YOU SMARTER THAN JUSTIN BIEBER?' is written in large, bold, red letters with a black outline. Below this, a photograph of Justin Bieber is shown on the right side, smiling and making a peace sign. On the left side, there is text: 'His IQ is above 102! Most People CANNOT beat him. Can you beat him?'. Below the text is a blue button with the Facebook logo and the text 'Connect with Facebook'. Underneath the button, there is a small line of text: '👍 37,527 people like this. Be the first of your friends to like this.'. At the bottom of the graphic, there is a red triangle logo and the text 'Are You Smart Enough!?' in white with a black outline. At the very bottom, there is a black bar with green text that reads 'Connect with your Facebook Account to see if you're SMARTER than Justin Bieber?'.

West, A. G. et al. (2011). Link spamming Wikipedia for profit. In *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference* (pp. 152-161). ACM.

WIKIMANIA '06

IDENTIFYING AND ELIMINATING TROLLS

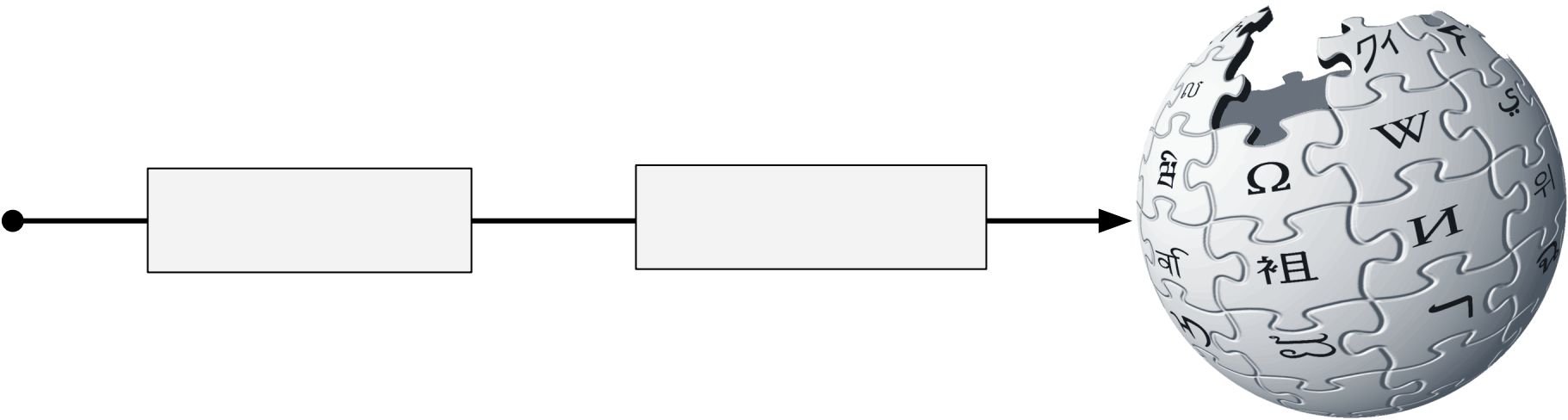




<https://commons.wikimedia.org/wiki/File:Bubbler.jpg>



https://commons.wikimedia.org/wiki/File:Flickr_-_Official_U.S._Navy_Imagery_-_Sailor%27s_daughter_operates_a_fire_hose_with_crew_member_assistance.jpg



Robots

Cyborgs



Machine learning

- Fast (~ 5 seconds)
- No human effort
- *Only obvious vandalism*

Human computation

- Still pretty fast (~ 30 seconds)
- Humans catch most vandalism at a glance

Huggle's UI

The screenshot displays the Huggle web interface. At the top, there is a navigation bar with menus for System, Queue, Go to, Revision, Page, User, Browser, and Help. A status bar on the right indicates '92 edits per minute, 4 reverts per minute'. Below the navigation bar is a toolbar with various icons for actions like undo, redo, and search. The main content area is titled 'Sehnsa' and shows two revisions. The left revision is from April 7, 2011, and the right revision is from April 14, 2011. The text of the revisions is identical: 'Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.' To the right of the text is a metadata table with fields like official_name, name, settlement_type, image_size, map_size, etc. A sidebar on the left lists 43 filtered edits, with 'Talk:Secession in the ...' highlighted. At the bottom, there is a chat log showing connection status to an IRC recent changes feed.

System Queue Go to Revision Page User Browser Help 92 edits per minute, 4 reverts per minute

Page Sehnsa History
User 91.109.230.97 Contribs

Sehnsa

Filtered edits
43 items
Kauffmann olefination
MicroBee
The Brothers Karamaz...
Barbajuan
Robert Conquest
Portuguese India Arma...
Chengdu Blades F.C.
Ryan Giggs
Child camel jockey
Raheem Sterling
Urza's Destiny
Talk:Secession in the ...
Gasoline and diesel us...
Little Things (short stor...
Steve Jobs
Saint Patapios
Alfred Fyodorov
Stefan Dennis
Labrador
Saeed al-Ghamdi
CS Obercorn
Ben Gurion Internation...
Bernstein's theorem on ...

Revision as of 16:01, April 7, 2011 (edit)
SmackBot (talk | contribs)
m (Dated {{Citation needed}}. (Build p608))

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

Latest revision as of 10:49, April 14, 2011 (edit)
91.109.230.97 (talk)
(→External links)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

{{unreferenced date=February 2009}}
{{Infobox settlement
official_name = Sehnsa
name =
settlement_type =
image_skyline =
image_size = 240px
image_alt =
image_caption =
image_map =
mapsize = 100px
map_alt =
map_caption =
latd = latm = lats = latNS =
longd = longm = longs = longEW =
coordinates_type = region:PK_type:city
coordinates_display =
pushpin_map = <!-- Pakistan -->
pushpin_label_position = <!-- left, right, top, bottom, none -->

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed
2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

Huggle's UI

The screenshot shows the Huggle web interface. At the top, there is a navigation bar with menus for System, Queue, Go to, Revision, Page, User, Browser, and Help. A status bar on the right indicates '92 edits per minute, 4 reverts per minute'. Below the navigation bar is a toolbar with various icons for editing and navigation. The main content area is titled 'Sehnsa' and shows two revisions. The left revision is from April 7, 2011, and the right revision is from April 14, 2011. Both revisions describe Sehnsa as a trading center in Pakistan. A photograph of a woman with a camera is visible in the center. On the left side, there is a sidebar with a 'Filtered edits' list containing 43 items, with 'Talk:Secession in the ...' highlighted. On the right side, there is a table of metadata for the selected revision, including fields like 'official_name', 'name', 'settlement_type', 'image_size', 'map_size', 'coordinates_type', etc. At the bottom, there are two status messages: '2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed' and '2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...'


System Queue Go to Revision Page User Browser Help 92 edits per minute, 4 reverts per minute

Page Sehnsa History
User 91.109.230.97 Contribs

Sehnsa

Revision as of 16:01, April 7, 2011 (edit)
SmackBot (talk | contribs)
m (Dated {{Citation needed}}). (Build p608)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.



Latest revision as of 10:49, April 14, 2011 (edit)
91.109.230.97 (talk)
(→External links)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

{{unreferenced date=February 2009}}
{{Infobox settlement
official_name = Sehnsa
name =
settlement_type =
image_skyline =
imagesize = 240px
image_alt =
image_caption =
image_map =
mapsize = 100px
map_at =
map_caption =
latd = latm = lats = latNS =
longd= longm= longs= longEW=
coordinates_type = region:PK_type:city
coordinates_display =
pushpin_map = <!-- Pakistan -->
pushpin_label_position = <!-- left, right, top, bottom, none -->

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed
2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

Huggle's UI

System Queue Go to Revision Page User Browser Help 92 edits per minute, 4 reverts per minute

Page: Sehnsa History

User: 91.109.230.97 Contribs

Filtered edits (43 items)

- Kauffmann olefination
- MicroBee
- The Brothers Karamaz...
- Barbajuan
- Robert Conquest
- Portuguese India Arma...
- Chengdu Blades F.C.
- Ryan Giggs
- Child camel jockey
- Raheem Sterling
- Urza's Destiny
- Talk:Secession in the ...**
- Gasoline and diesel us...
- Little Things (short stor...
- Steve Jobs
- Saint Patapios
- Alfred Fyodorov
- Stefan Dennis
- Labrador
- Saeed al-Ghamdi
- CS Obercorn
- Ben Gurion Internation...
- Bernstein's theorem on ...

Sehnsa

Revision as of 10:49, April 14, 2011 (edit)

91.109.230.97 (talk) (→External links)

Sehnsa is a city in the Punjab province of Pakistan, which is one of the two main cities of Pakistan, which are Islamabad and Kotli.

Sehnsa city is also well known trading centre because of its easy access to two main cities of Pakistan, which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

```
{{unreferenced|date=February 2009}}
{{Infobox settlement
|official_name = Sehnsa
|name =
|settlement_type =
|image_skyline =
|imagesize = 240px
|image_alt =
|image_caption =
|image_map =
|mapsize = 100px
|map_at =
|map_caption =
|latd = | latm = | lats = | latNS =
|longd = | longm = | longs = | longEW=
|coordinates_type = region:PK_type:city
|coordinates_display =
|pushpin_map = <!-- Pakistan -->
|pushpin_label_position = <!-- left, right, top, bottom, none -->
```

+6128

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed
2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

Huggle's UI

The screenshot displays the Huggle web interface. At the top, there is a navigation bar with menus for System, Queue, Go to, Revision, Page, User, Browser, and Help. A status bar on the right indicates '92 edits per minute, 4 reverts per minute'. Below the navigation bar is a toolbar with various icons for actions like search, edit, and revert. The main content area is titled 'Sehnsa' and shows two revision entries. The left entry is from April 7, 2011, and the right entry is from April 14, 2011. The text of the revisions is identical: 'Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.' To the right of the text is a metadata table with fields like official_name, name, settlement_type, image_size, map_size, etc. On the left side, there is a sidebar with a 'Filtered edits' dropdown and a list of 43 items, including 'Kauffmann olefination', 'MicroBee', 'The Brothers Karamaz...', 'Barbajuan', 'Robert Conquest', 'Portuguese India Arma...', 'Chengdu Blades F.C.', 'Ryan Giggs', 'Child camel jockey', 'Raheem Sterling', 'Urza's Destiny', 'Talk:Secession in the ...', 'Gasoline and diesel us...', 'Little Things (short stor...', 'Steve Jobs', 'Saint Patapios', 'Alfred Fyodorov', 'Stefan Dennis', 'Labrador', 'Saeed al-Ghamdi', 'CS Obercorn', 'Ben Gurion Internation...', and 'Bernstein's theorem on ...'. At the bottom, there is a chat log showing IRC messages.

System Queue Go to Revision Page User Browser Help 92 edits per minute, 4 reverts per minute

Page Sehnsa History
User 91.109.230.97 Contribs

Sehnsa

Revision as of 16:01, April 7, 2011 (edit)
SmackBot (talk | contribs)
m (Dated {{Citation needed}}). (Build p608)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

Latest revision as of 10:49, April 14, 2011 (edit)
91.109.230.97 (talk)
(→External links)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

{{unreferenced date=February 2009}}
{{Infobox settlement
official_name = Sehnsa
name =
settlement_type =
image_skyline =
imagesize = 240px
image_alt =
image_caption =
image_map =
mapsize = 100px
map_alt =
map_caption =
latd = latm = lats = latNS =
longd = longm = longs = longEW =
coordinates_type = region:PK_type:city
coordinates_display =
pushpin_map = <!-- Pakistan -->
pushpin_label_position = <!-- left, right, top, bottom, none -->

Filtered edits
43 items
Kauffmann olefination
MicroBee
The Brothers Karamaz...
Barbajuan
Robert Conquest
Portuguese India Arma...
Chengdu Blades F.C.
Ryan Giggs
Child camel jockey
Raheem Sterling
Urza's Destiny
Talk:Secession in the ...
Gasoline and diesel us...
Little Things (short stor...
Steve Jobs
Saint Patapios
Alfred Fyodorov
Stefan Dennis
Labrador
Saeed al-Ghamdi
CS Obercorn
Ben Gurion Internation...
Bernstein's theorem on ...

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed
2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

Huggle's UI

The screenshot displays the Huggle web interface. At the top, there is a navigation bar with menus for System, Queue, Go to, Revision, Page, User, Browser, and Help. A status bar on the right indicates '92 edits per minute, 4 reverts per minute'. Below the navigation bar is a toolbar with various icons for actions like search, edit, and revert. The main content area is divided into two columns. The left column shows a list of 'Filtered edits' with 43 items, including 'Kauffmann olefination', 'MicroBee', 'The Brothers Karamaz...', 'Barbajuan', 'Robert Conquest', 'Portuguese India Arma...', 'Chengdu Blades F.C.', 'Ryan Giggs', 'Child camel jockey', 'Raheem Sterling', 'Urza's Destiny', 'Talk:Secession in the ...', 'Gasoline and diesel us...', 'Little Things (short stor...', 'Steve Jobs', 'Saint Patapios', 'Alfred Fyodorov', 'Stefan Dennis', 'Labrador', 'Saeed al-Ghamdi', 'CS Obercorn', 'Ben Gurion Internation...', and 'Bernstein's theorem on ...'. The right column shows the search results for 'Sehnsa'. It includes a 'Revision as of 16:01, April 7, 2011 (edit)' by SmackBot and a 'Latest revision as of 10:49, April 14, 2011 (edit)' by user 91.109.230.97. The text of the article is: 'Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.' Below the text is a table of metadata for the article, including fields like 'official_name', 'name', 'settlement_type', 'image_size', 'map_size', 'coordinates_type', etc. At the bottom of the interface, there is a chat log showing messages from 2011-04-14 10:48:55 and 2011-04-14 10:48:48.

Huggle

System Queue Go to Revision Page User Browser Help

92 edits per minute, 4 reverts per minute

Page Sehnsa History

User 91.109.230.97 Contribs

Sehnsa

Revision as of 16:01, April 7, 2011 (edit)
SmackBot (talk | contribs)
m (Dated {{Citation needed}}). (Build p608)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

Latest revision as of 10:49, April 14, 2011 (edit)
91.109.230.97 (talk)
(→External links)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

{{unreferenced date=February 2009}}
{{Infobox settlement
official_name = Sehnsa
name =
settlement_type =
image_skyline =
imagesize = 240px
image_alt =
image_caption =
image_map =
mapsize = 100px
map_at =
map_caption =
latd = latm = lats = latNS =
longd = longm = longs = longEW =
coordinates_type = region:PK_type:city
coordinates_display =
pushpin_map = <!-- Pakistan -->
pushpin_label_position = <!-- left, right, top, bottom, none -->

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed

2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

The Worldview of Huggle

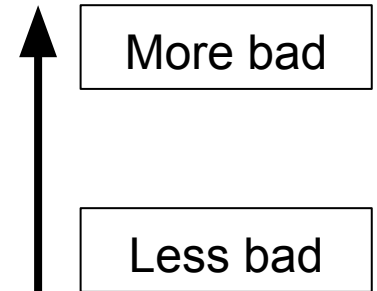
The Worldview of Huggle

- Wikipedia is a firehose of contributions that need **review**.



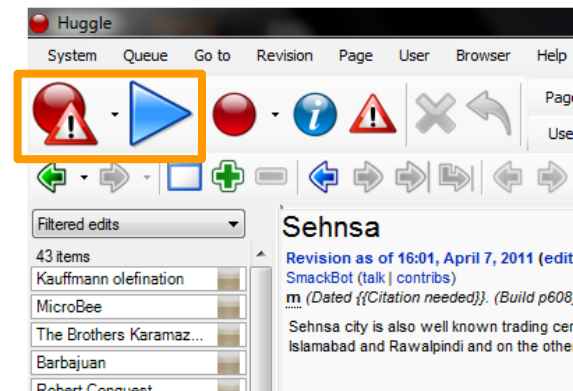
The Worldview of Huggle

- Wikipedia is a firehose of contributions that need **review**.
- Use algorithms to **triage** & deal with most suspicious first



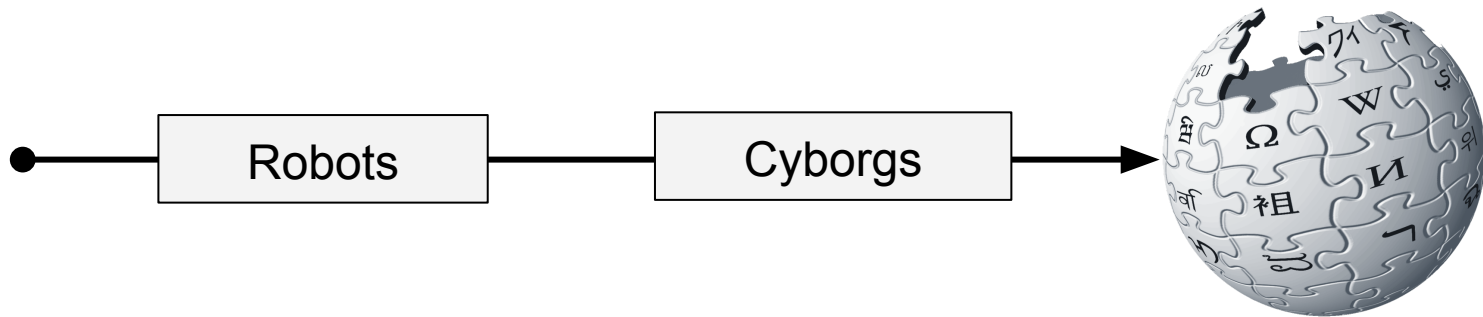
The Worldview of Huggle

- Wikipedia is a firehose of contributions that need **review**.
- Use algorithms to **triage** & deal with most suspicious first
- **Reject** contributions that don't meet standards.



Wild Success!

- Saved Wikipedia from vandalism[1,3]
- Dramatically reduced Wikipedian effort[2]



1. Priedhorsky, R. et al. (2007). Creating, destroying, and restoring value in Wikipedia. GROUP (pp. 259-268). ACM.
2. Halfaker, A. Vandal Fighter Work Load. (2011). https://meta.wikimedia.org/wiki/Research:Vandal_fighter_work_load
3. Geiger, R. S., & Halfaker, A. (2013). When the levee breaks: without bots, what happens to Wikipedia's quality control processes? *WikiSym* (p. 6). ACM.

But...

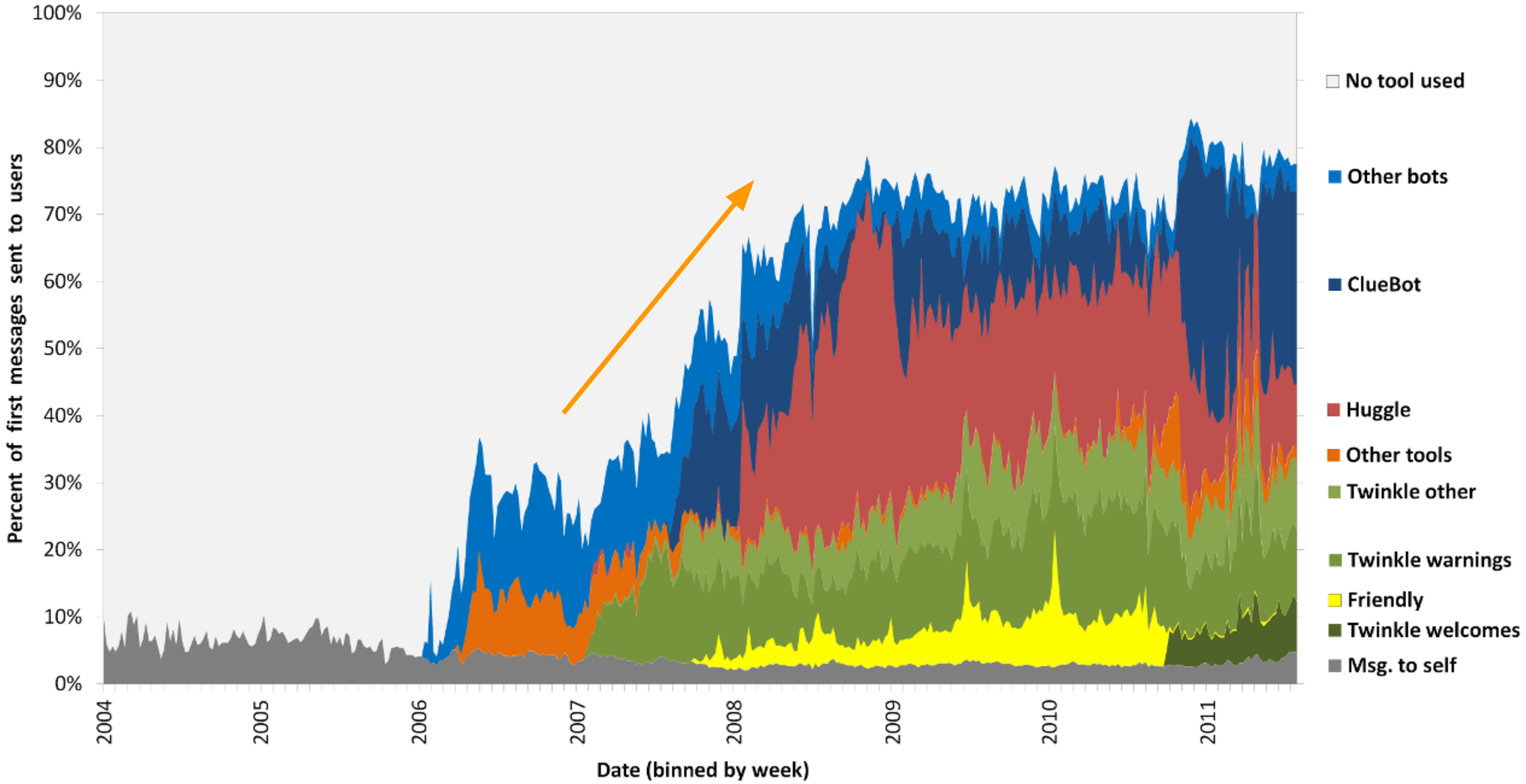
There's some unforeseen consequences

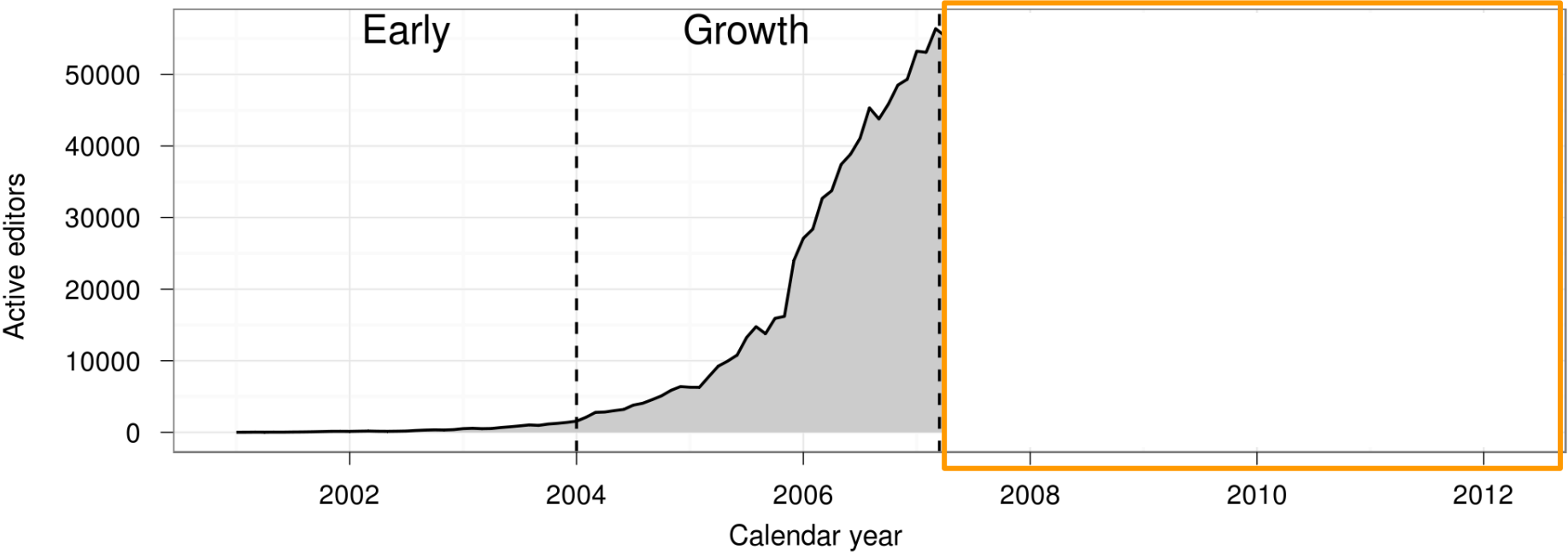


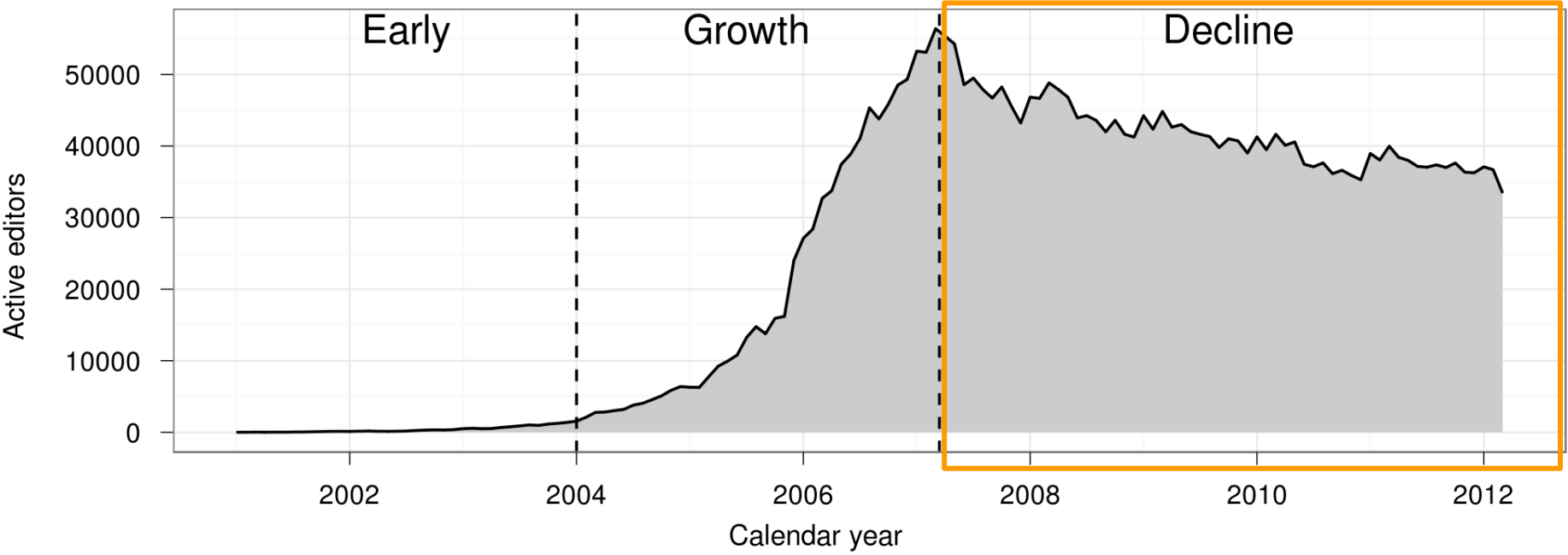
Reverts

Warnings

First messages sent to new users, by tool and bot used (proportional)

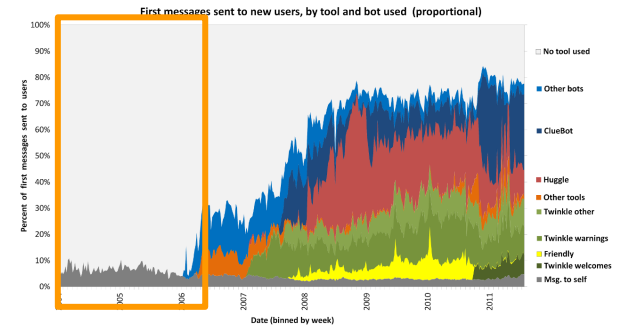






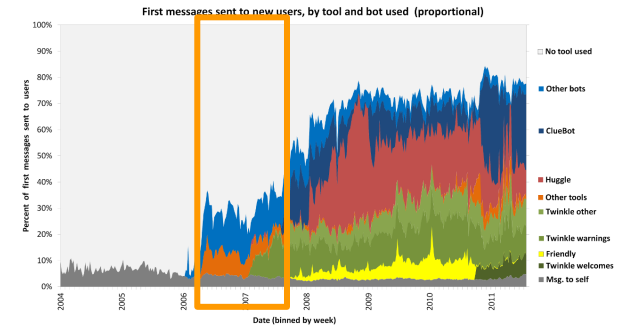
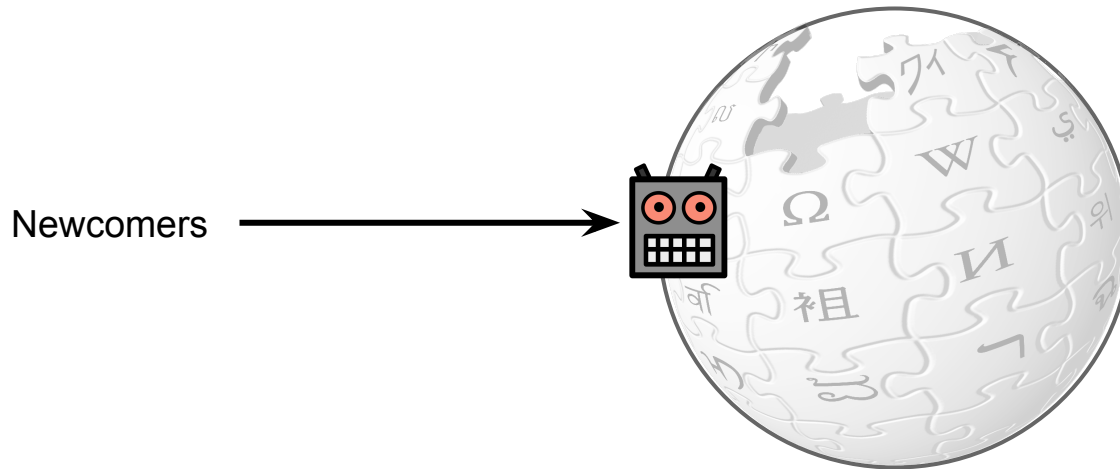
What happened?

Newcomers



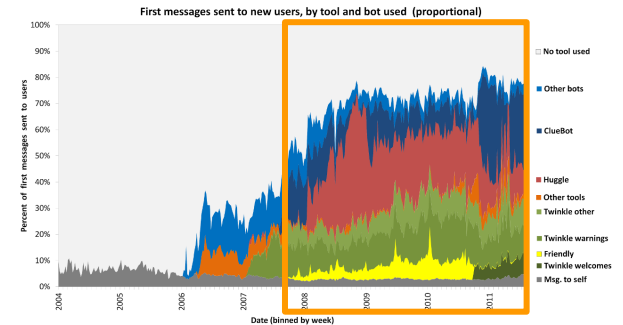
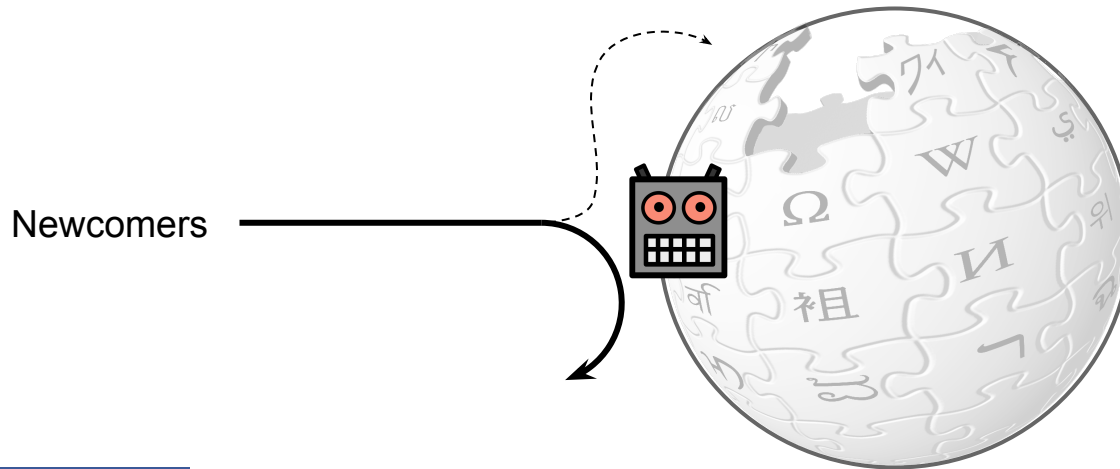
1. Halfaker, A., Gieger, R. S., Morgan, J., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System: *American Behavioral Scientist* 57(5) 664-688.

What happened?



1. Halfaker, A., Gieger, R. S., Morgan, J., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System: *American Behavioral Scientist* 57(5) 664-688.

What happened?



1. Halfaker, A., Gieger, R. S., Morgan, J., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System. *American Behavioral Scientist* 57(5) 664-688.

What can be done?

What can be done?

- Efficient quality control is necessary

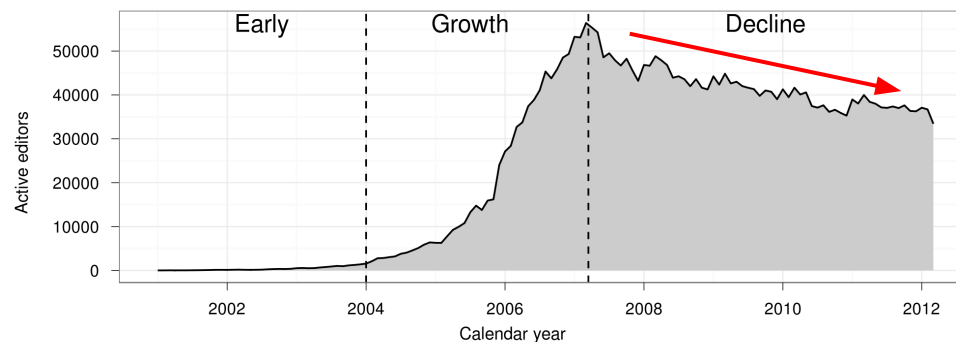


What can be done?

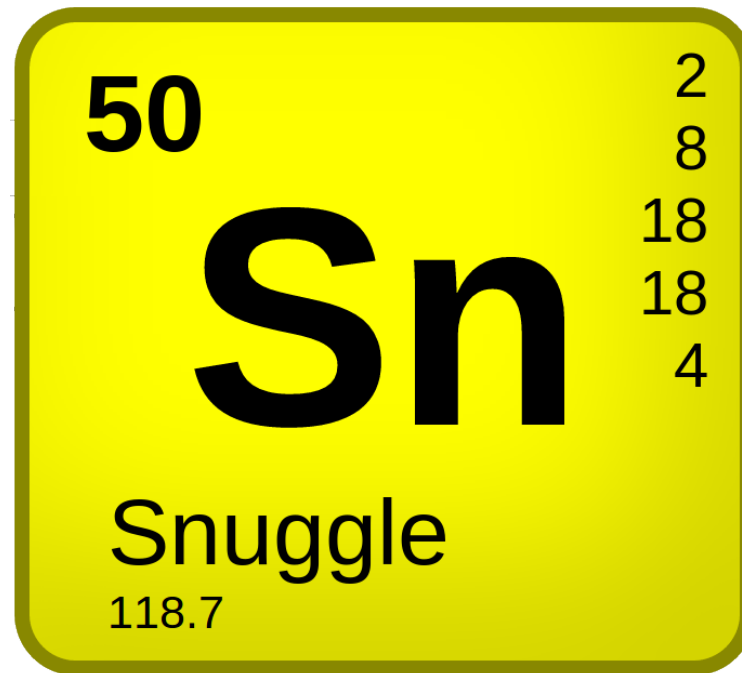
- Efficient quality control is necessary



- Socializing new editors is necessary



Part 2



Part 2



Newcomer
socialization

vs.



Counter-
vandalism

Intelligent algorithms for triage

The screenshot shows the Huggle interface with a menu bar (System, Queue, Go to, Revision, Page, User, Browser, Help) and a status bar (92 edits per minute, 4 reverts per minute). The main area is divided into a left sidebar for 'Filtered edits' and a main content area for the 'Sehnsa' page. The 'Filtered edits' list includes items like 'Kauffmann olefination', 'MicroBee', and 'Talk:Secession in the ...'. The 'Sehnsa' page shows the latest revision as of 10:49, April 14, 2011, with a description of the city and its location. A black arrow points from the 'Talk:Secession in the ...' item in the list to the 'Sehnsa' page. Two text boxes are overlaid on the image: 'More bad' with an arrow pointing to the 'Talk:Secession in the ...' item, and 'Less bad' with an arrow pointing to the 'Sehnsa' page.

More bad

Less bad

Intelligent algorithms for triage

The screenshot shows the Snuggle website interface. The browser address bar displays <https://snuggle.grouplens.org>. The user is not logged in. The main navigation bar includes tabs for **uncategorized**, **good-faith**, **ambiguous**, and **bad-faith**. Below this, a filter bar shows **uncategorized** selected, with options for newcomers with more than 1 edit to any page, sorted by desirability in descending order.

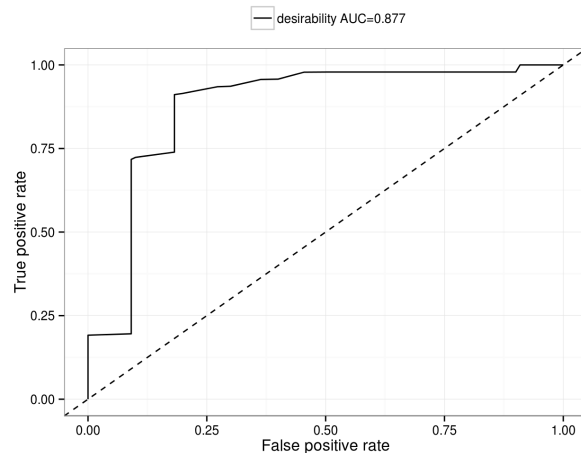
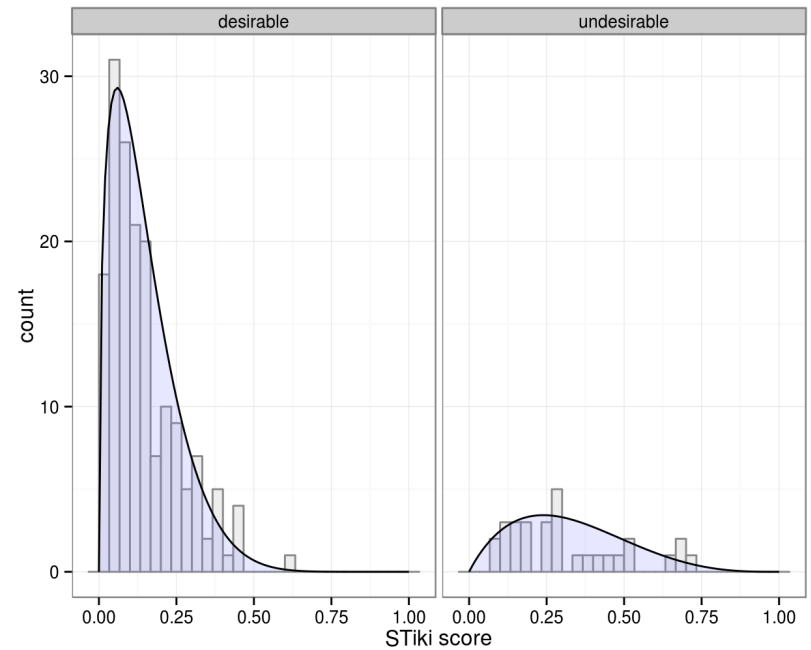
Username and stats	Interactive edit graph	Talk page activity
<input type="checkbox"/> Sinineburx		<input type="radio"/>
<input type="checkbox"/> Betabot75		<input type="radio"/> W warn
<input type="checkbox"/> Rodriguezandres999		<input type="radio"/> vand¹ TH
<input type="checkbox"/> PrisciliaFria		<input type="radio"/> csd <input type="radio"/>
<input type="checkbox"/> RichardHC		<input type="radio"/> TH
<input type="checkbox"/> J3sseung		<input type="radio"/>
<input type="checkbox"/> Lostbad		<input type="radio"/> W vand¹ TH
<input type="checkbox"/> Wikihj4		
<input type="checkbox"/> Dswatcha		<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>
<input type="checkbox"/> Keaclamvietot		
<input type="checkbox"/> Meleri Morgan		<input type="radio"/> W TH
<input type="checkbox"/> Tomski12		<input type="radio"/> vand¹
<input type="checkbox"/> Nordicanuck		<input type="radio"/> W
<input type="checkbox"/> MarielMarshall		<input type="radio"/> <input type="radio"/>

Intelligent algorithms for triage

The screenshot shows the Snuggle website interface. At the top, there are navigation tabs: 'uncategorized', 'good-faith', 'ambiguous', and 'bad-faith'. The 'good-faith' tab is selected. Below the tabs, there are filters for 'newcomers with more than 1 edits to any page' and sorting options 'desirability' (ascending/descending). The main content area is a table with three columns: 'Username and stats', 'Interactive edit graph', and 'Talk page activity'. The table lists several users, each with a checkbox, a bar chart representing edit activity, and various activity indicators (e.g., 'W', 'warn', 'vand¹', 'TH', 'csd'). A vertical black arrow points upwards from the bottom of the table to the top, indicating a sorting algorithm. Two text boxes are overlaid on the arrow: 'More good' at the top and 'Less good' at the bottom.

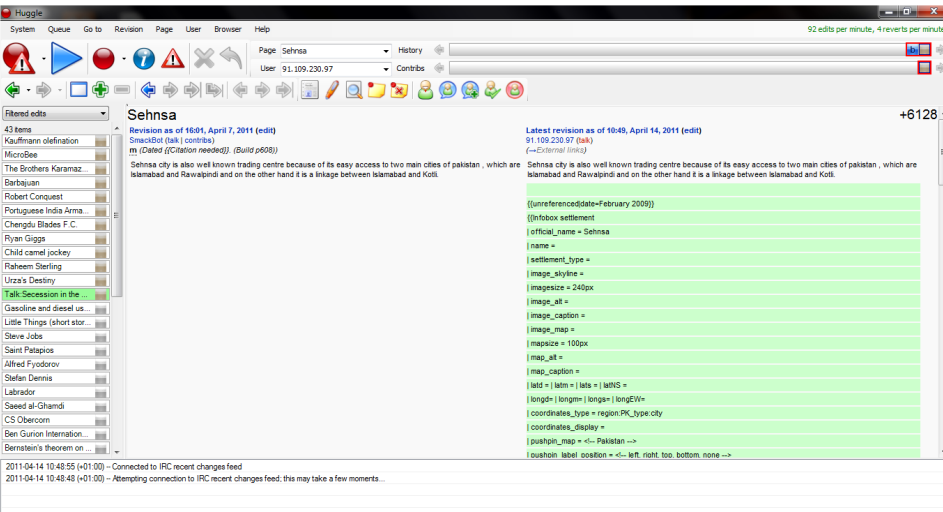
Username and stats	Interactive edit graph	Talk page activity
<input type="checkbox"/> Sinineburx	[Bar chart]	
<input type="checkbox"/> Betabot75	[Bar chart]	W warn
<input type="checkbox"/> Rodriguezandres999	[Bar chart]	vand¹ TH
<input type="checkbox"/> PrisciliaFrilia	[Bar chart]	csd
<input type="checkbox"/> RichardHC	[Bar chart]	TH
<input type="checkbox"/> J3sseung	[Bar chart]	
<input type="checkbox"/> Lostbad	[Bar chart]	W vand¹ TH
<input type="checkbox"/> Wikihj4	[Bar chart]	
<input type="checkbox"/> Dswatcha	[Bar chart]	
<input type="checkbox"/> Keaclamvietot	[Bar chart]	
<input type="checkbox"/> Meleri Morgan	[Bar chart]	W TH
<input type="checkbox"/> Tomski12	[Bar chart]	vand¹
<input type="checkbox"/> Nordicanuck	[Bar chart]	W
<input type="checkbox"/> MarielMarshall	[Bar chart]	

Intelligent algorithms for triage



**1st day's edits:
0.877 AUC**

Social literacy

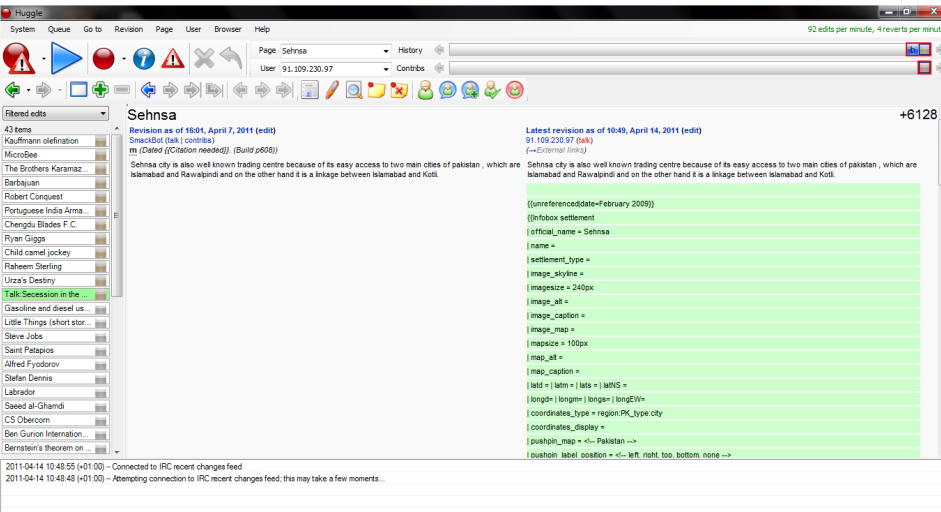


Warning template levels:

- 1
- 2
- 3
- 4
- → Notify the admins!

1. Geiger, R. S., & Ribes, D. (2010). The work of sustaining order in wikipedia. CSCW (pp. 117-126). ACM.

Social literacy



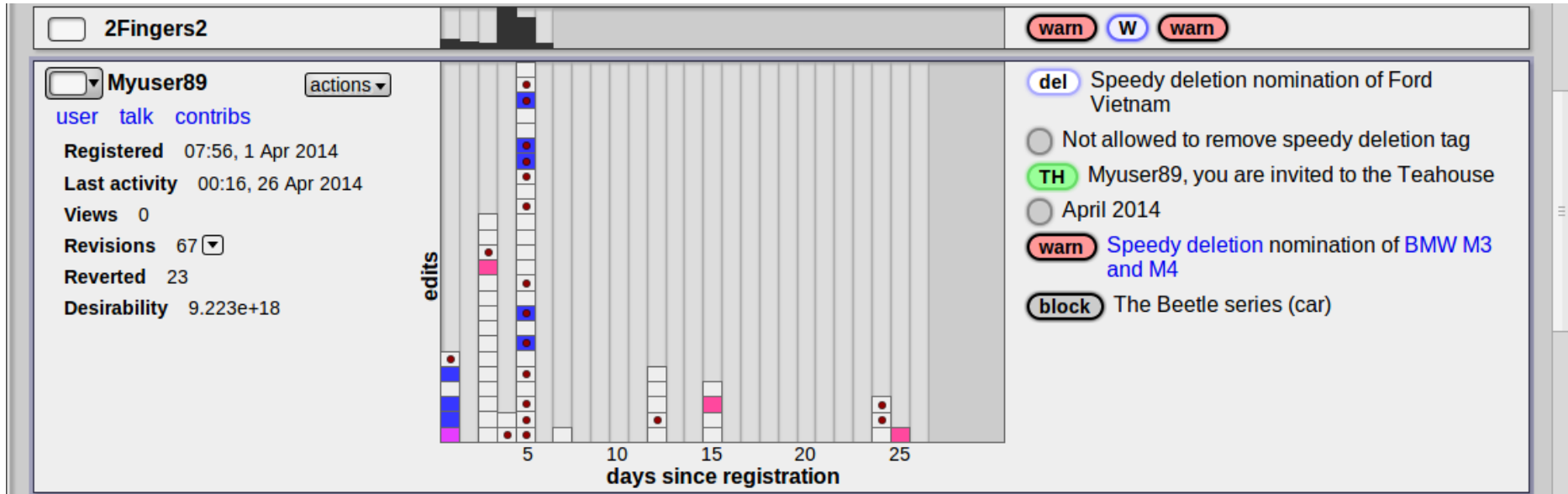
Warning template levels:

- 1
- 2
- 3
- 4
- → Notify the admins!

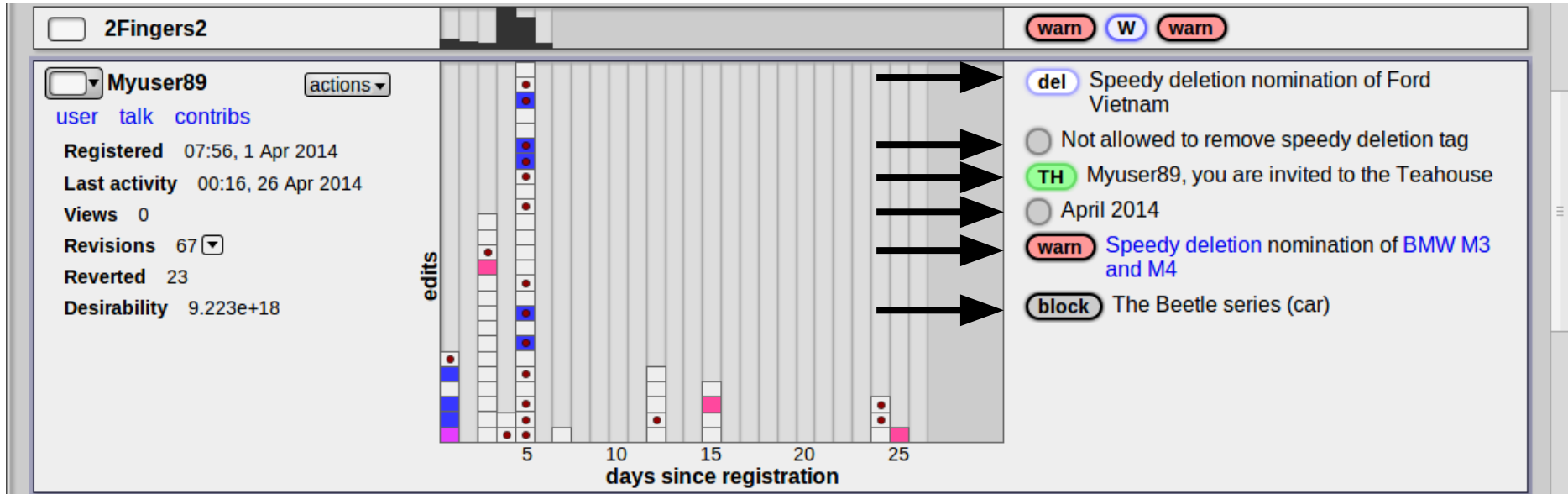
Transparent to the user.

1. Geiger, R. S., & Ribes, D. (2010). The work of sustaining order in wikipedia. CSCW (pp. 117-126). ACM.

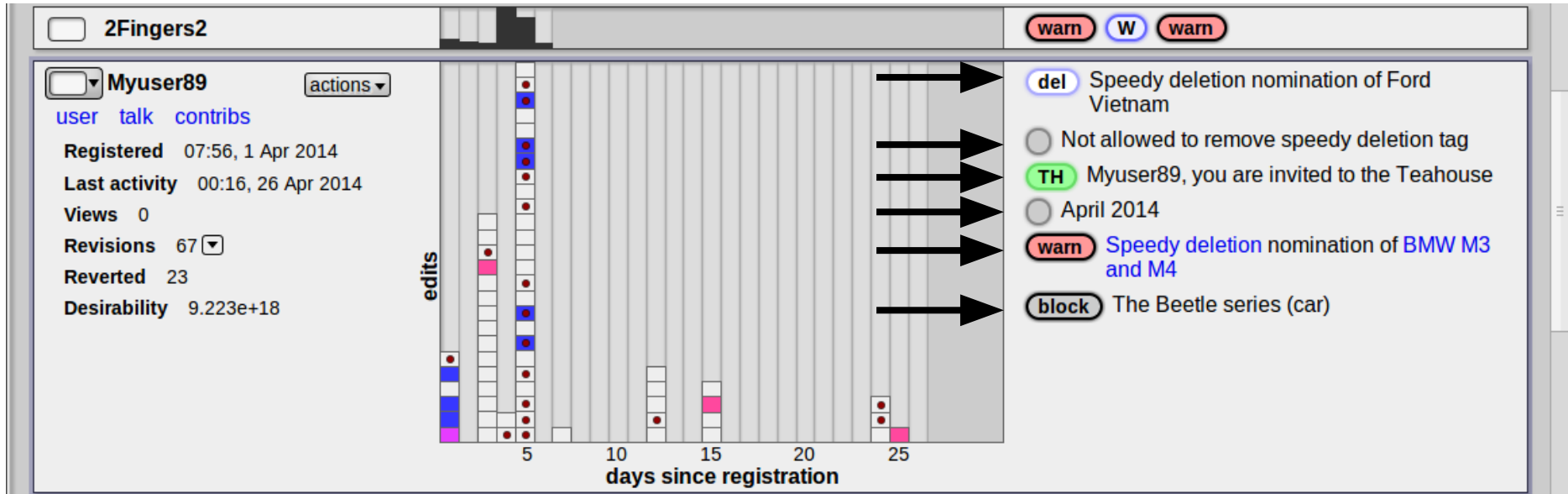
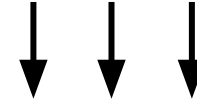
Social literacy



Social literacy



Social literacy



Social literacy

2Fingers2

warn W warn

Myuser89 actions del Speedy deletion nomination of Ford

user talk contribute send message

Registered 07: invite to teahouse Posts an invitation to the WP:Teahouse on Myuser89's talk page.

Last activity 0 report abuse

Views 0

Revisions 67

Reverted 23

Desirability 9.223e+18

Template

Invitation Invitation2

Header

You're invited to the [[WP:Teahouse|Teah

Message

Append to: User_talk:Myuser89

You're invited to the [Teahouse](#).^[edit]

Hello **Myuser89**, thanks for contributing to Wikipedia! I want to invite you to the [Teahouse](#), a forum on Wikipedia for new editors to ask questions about editing Wikipedia, and get support from peers and experienced editors. I hope you see you there! [EpochFail](#) (talk • contribs) 22:18, 27 April 2014 (UTC)

summary: You're invited to the [Teahouse](#). (Sn)

preview

watch user cancel submit

Holistic picture of another human

Holistic picture of another human

The screenshot shows the Huggle web interface. At the top, there's a navigation bar with 'System', 'Queue', 'Go to', 'Revision', 'Page', 'User', 'Browser', and 'Help'. The page title is 'Sehnsa' and the user is '91.109.230.97'. The interface includes a toolbar with various icons for navigation and editing. On the left, there's a 'Filtered edits' sidebar with a list of 43 items, including 'Kauffmann olefination', 'MicroBee', and 'Talk:Secession in the ...'. The main content area displays the article for 'Sehnsa', showing two revisions: one from April 7, 2011, and a latest revision from April 14, 2011. The article text describes Sehnsa as a trading center in Pakistan. To the right of the article, there's a table of metadata for the settlement, including fields like 'official_name', 'name', 'settlement_type', 'image_size', 'map_size', and 'coordinates_type'. The bottom of the interface shows a status bar with connection logs for IRC recent changes feed.

System Queue Go to Revision Page User Browser Help 92 edits per minute, 4 reverts per minute

Page Sehnsa History
User 91.109.230.97 Contribs

Filtered edits
43 items
Kauffmann olefination
MicroBee
The Brothers Karamaz...
Barbajuan
Robert Conquest
Portuguese India Arma...
Chengdu Blades F.C.
Ryan Giggs
Child camel jockey
Raheem Sterling
Urza's Destiny
Talk:Secession in the ...
Gasoline and diesel us...
Little Things (short stor...
Steve Jobs
Saint Patapios
Alfred Fyodorov
Stefan Dennis
Labrador
Saeed al-Ghamdi
CS Obercorn
Ben Gurion Internation...
Bernstein's theorem on ...

Sehnsa

Revision as of 16:01, April 7, 2011 (edit)
SmackBot (talk | contribs)
m (Dated {{Citation needed}}). (Build p608)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

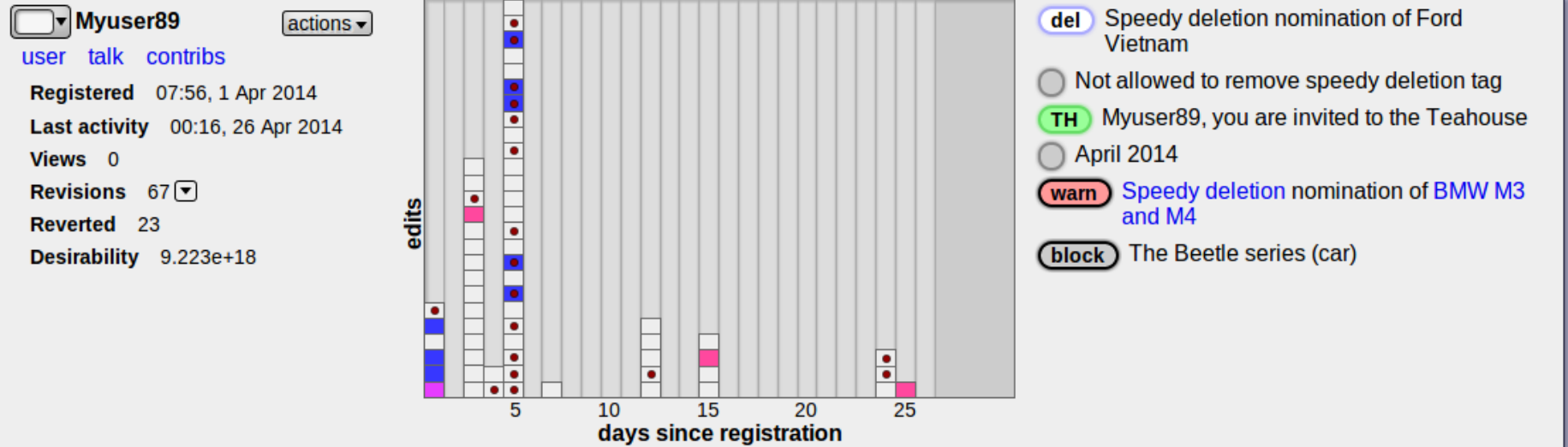
Latest revision as of 10:49, April 14, 2011 (edit)
91.109.230.97 (talk)
(→External links)

Sehnsa city is also well known trading centre because of its easy access to two main cities of pakistan , which are Islamabad and Rawalpindi and on the other hand it is a linkage between Islamabad and Kotli.

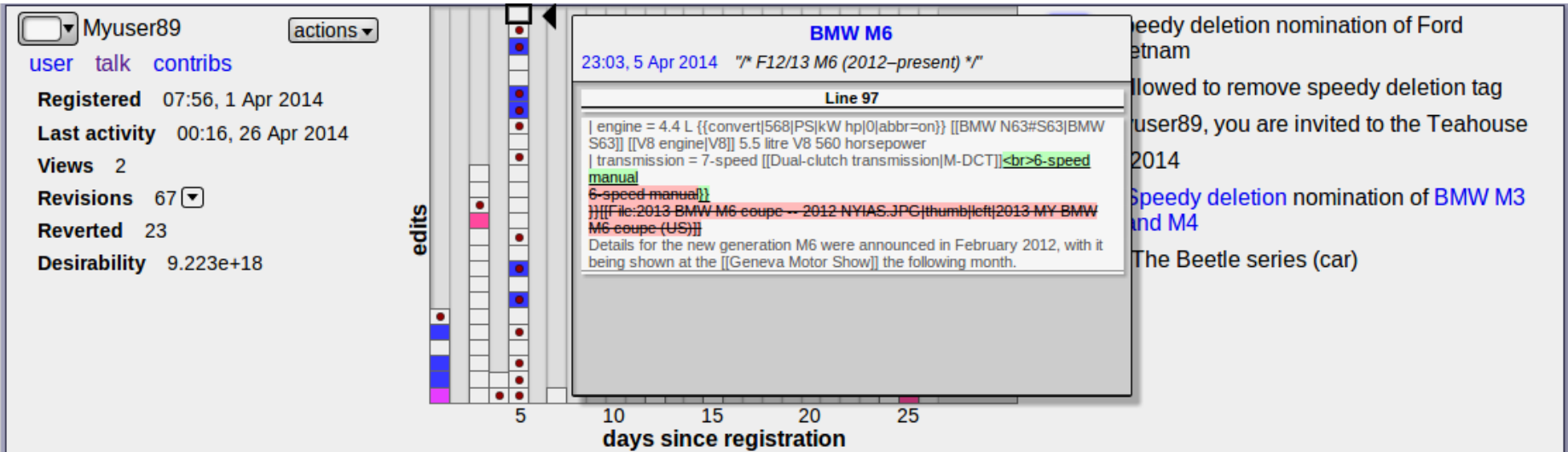
{{unreferenced date=February 2009}}
{{Infobox settlement
official_name = Sehnsa
name =
settlement_type =
image_skyline =
imagesize = 240px
image_alt =
image_caption =
image_map =
mapsize = 100px
map_alt =
map_caption =
latd = latm = lats = latNS =
longd= longm= longs= longEW=
coordinates_type = region:PK_type:city
coordinates_display =
pushpin_map = <!-- Pakistan -->
oushoin_label_position = <!-- left, right, top, bottom, none -->

2011-04-14 10:48:55 (+01:00) -- Connected to IRC recent changes feed
2011-04-14 10:48:48 (+01:00) -- Attempting connection to IRC recent changes feed; this may take a few moments...

Holistic picture of another human



Holistic picture of another human



Holistic picture of another human

Myuser89 actions

[user](#) [talk](#) [contribs](#)

Registered 07:56, 1 Apr 2014

Last activity 00:16, 26 Apr 2014

Views 2

Revisions 67 ▾

Reverted 23

Desirability 9.223e+18

Chevrolet Volt

22:14, 5 Apr 2014 *(no edit summary)*

- Reverted by Mr.choppers "Undid revision 602948131 by Myuser89 (talk)"

Line 640

```
{{Chevrolet cars timeline 1980 to date}}
{{Navboxes|list1=
{{Opel}}
{{Opel timeline}}
{{Vauxhall Motors}}
}}
```

Speedy deletion nomination of Ford
etnam

Allowed to remove speedy deletion tag
user89, you are invited to the Teahouse
2014

Speedy deletion nomination of BMW M3
and M4

The Beetle series (car)

Holistic picture of another human

Myuser89 actions

[user](#) [talk](#) [contribs](#)

Registered 07:56, 1 Apr 2014

Last activity 00:16, 26 Apr 2014

Views 2

Revisions 67 ▾

Reverted 23

Desirability 9.223e+18

edits

Template:Opel timeline 1980 to present
22:12, 5 Apr 2014 (no edit summary)

- Reverted by Mr.choppers "Undid revision 602948001 by Myuser89 (talk) fix link to redirect"

Line 50

```
|colspan=1|
|colspan=5 bgcolor=#C0C0C0| [[ChevroletOpel Ve#Ampera (car)|Ampera]]
|- bgcolor=#E0E0E0
```

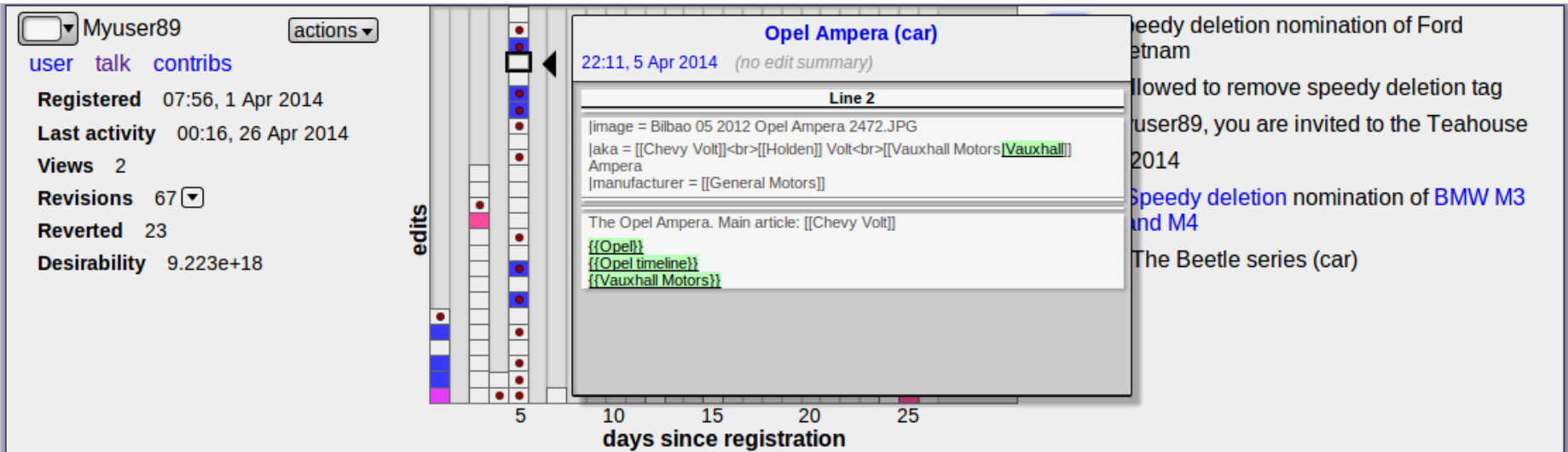
Speedy deletion nomination of Ford
etnam

Allowed to remove speedy deletion tag
user89, you are invited to the Teahouse
2014

Speedy deletion nomination of BMW M3
and M4

The Beetle series (car)

Holistic picture of another human



Draw attention to the social problem

Draw attention to the social problem



Draw attention to the social problem



This close-up shows a user's talk page activity. At the top, there are several action buttons: W, vand², and csd. Below these are two rows of buttons: W, TH and W, TH. The main content area shows a list of actions with arrows pointing to them: a "warn" button, a "del" button followed by the text "Speedy deletion nomination of Ford Vietnam", a "TH" button followed by "Myuser89, you are invited to the Teahouse", a "warn" button followed by "Speedy deletion nomination of BMW M3 and M4", and a "block" button followed by "The Beetle series (car)".

Draw attention to the social problem

Interviewee #10:

Draw attention to the social problem

Interviewee #10:

I don't see why this guy was reverted.

Draw attention to the social problem

Interviewee #10:

I don't see why this guy was reverted.

I don't see how this is vandalism.

Draw attention to the social problem

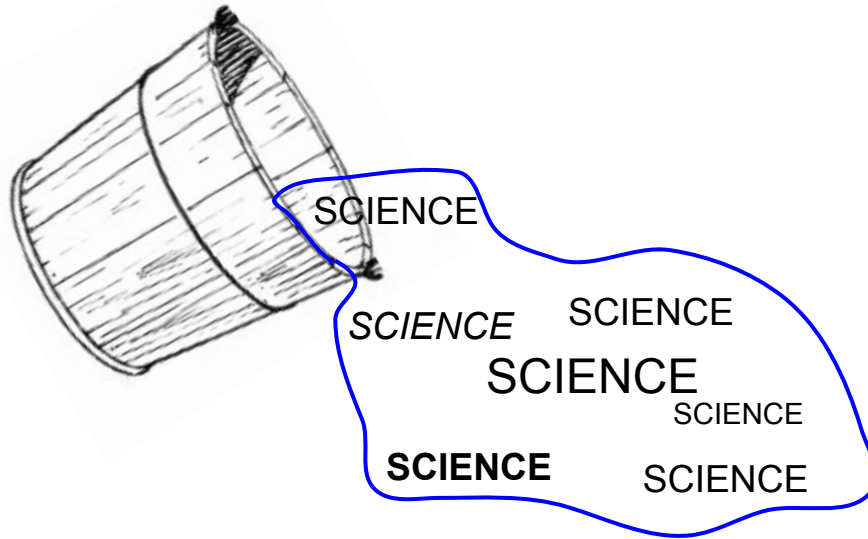
Interviewee #10:

I don't see why this guy was reverted.

I don't see how this is vandalism.

This is a false positive.

Thanks!



NSF Funding

- IIS 11-11201
- IIS 08-08692

My collaborators

- R. Stuart Geiger
- Loren Terveen

Aaron Halfaker

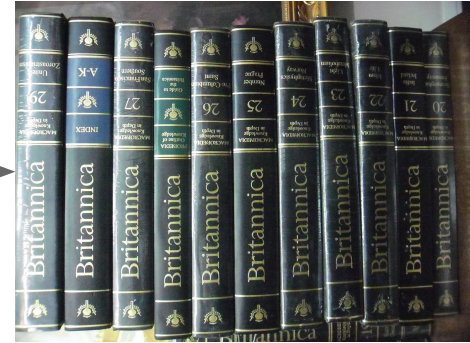
ahalfaker@wikimedia.org

Send me emails about science, Wikipedia, the internet or whatever.

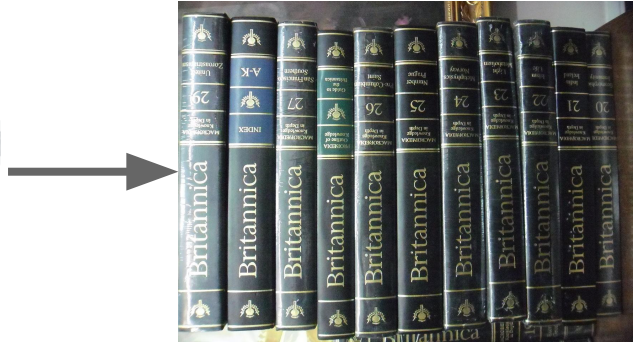
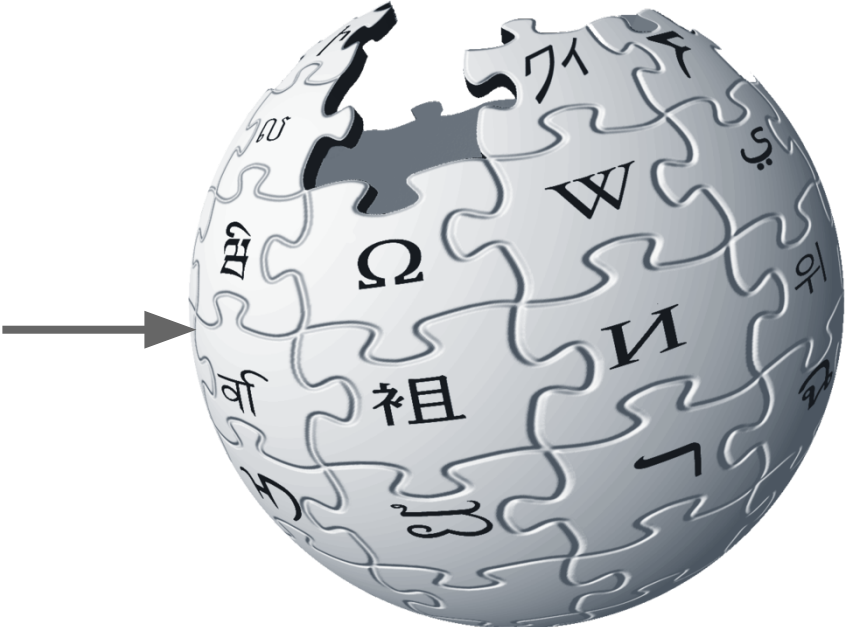
Ping me (halfak) in #wikimedia-research on freenode

<http://snuggle.grouplens.org>

Volunteers'
time & energy



Volunteers'
time & energy



~1:10

Planet Earth
(6.97 billion)

Unique
Readers
(500 million)

estimates from 2012

